

版权注意事项：

- 1、书籍版权归作者和出版社所有
- 2、本PDF仅限用于个人获取知识，进行私底下的知识交流
- 3、PDF获得者不得在互联网上以任何目的进行传播
- 4、如觉得书籍内容很赞，请购买正版实体书，支持作者
- 5、请于下载PDF后24小时内删除本PDF。

DRIVERLESS 无人驾驶

人工智能将从颠覆驾驶开始，全面重构人类生活
一场关于交通、物流、能源、制造、保险、医疗、伦理道德的全面挑战已经开始

【美】
胡迪·利普森
梅尔芭·库曼 著
(《3D打印》作者)
林露茵 金阳 译



Intelligent Cars and
the Road Ahead

by Hod Lipson and
Melba Kurman



无人驾驶

人工智能将从颠覆驾驶开始,全面重构人类生活

你初次拥有的智能机器人,很可能就是你的汽车,你会把自己的性命交给它!

● 伴随着无人驾驶技术的普及,人工智能将改变人类对空间的认知,重塑社会的结构:住所位置将进一步外延,城市面积将进一步扩大;市内停车场将被重新规划为住宅和商业空间;医疗和交通设施会随着车祸率降低而被推倒重建;人类的工作、购物、娱乐习惯将被全面颠覆,整个社会将面临人工智能的洗牌……但这只是人工智能革命的开始。

● 人工智能还将挑战人类传统伦理和道德的边际:全球数千万全职司机将面临失业;公共运输行业将会整体凋零;能源和制造行业将被迫转型;汽车产业和保险业将被彻底颠覆;人类可能需要牺牲个人隐私,换取出行的安全和便利性……在遭遇突发状况时,无人驾驶汽车将如何选择?

● 人类如何实现向人工智能时代的平稳过渡?哥伦比亚大学人工智能实验室主任胡迪·利普森教授将以科学、严谨、全面、系统的论述,揭开人工智能时代的序幕。



无人驾驶

人工智能将从颠覆驾驶开始，全面重构人类生活

〔美〕胡迪·利普森 梅尔芭·库曼 著

林露茵 金阳 译



DRIVERLESS

Intelligent Cars and the Road Ahead
by Hod Lipson and Melba Kurman

文匯出版社

目录

前言 / 001

导语 / 005

机器司机 / 007

迎接无人驾驶 / 016

致命病症的特效药 / 021

阻碍无人驾驶技术发展的7个误区 / 023

无人驾驶技术发展的时间表 / 027

第一章 无人驾驶的世界 / 031

便利的个人移动性 / 035

找停车位不再令人头疼 / 043

更短的通勤时间和更少花费 / 049

社交和孤独感 / 052

第二章 终极的移动设备 / 057

汽车和代码 / 061

大动荡 / 064

以人类为主导 / 070

第三章 独立思维 / 081

无人驾驶汽车的操作系统 / 085

机器人研发的艺术 / 086

运动中的人工智能 / 091

驾驶的控制权：混合人工智能 / 094

底层控制系统：加速、刹车和转向 / 097

上层控制系统：路径规划和道路导航 / 101

第四章 创建人工识别 / 105

物体识别的挑战 / 109

中层控制系统 / 116

绝对安全可靠 / 121

比人类驾驶安全两倍 / 126

第五章 前提——电子化的高速公路 / 131

黄金时代 / 139

通用汽车的电子化高速公路 / 143

自动化高速公路的衰亡 / 150

第六章 打造智能汽车，而非智能高速公路 / 155

V2X技术 / 158

智能交通系统的历史 / 161

关于V2X的反思 / 166

非智能高速公路的价值 / 172

修订交通运输政策 / 174

第七章 机器人的崛起 / 181

机器学习和无人驾驶 / 186

西洋跳棋 / 192

无限状态空间 / 198

现代工具箱 / 200

第八章 无人驾驶汽车全解读 / 207

高清数字地图 / 209

数码相机 / 211

光检测与测距（激光雷达） / 215

无线电检测与测距（雷达） / 219

超声波传感器（声呐） / 223

全球定位系统（GPS） / 223

内置的耳朵（IMU惯性测量单元） / 225

线控技术（Drive by Wire） / 229

其他设备 / 233

第九章 深度学习，无人驾驶的最后一块拼图 / 235

神经网络 / 239

感知器 / 242

第一次人工智能冬天 / 248

神经网络研究的复兴 / 251

神经认知机 / 255

深度学习的诞生 / 259

绘图处理器 (GPU_s) / 262

现代深度学习 / 264

神经网络内部 / 267

新型边缘检测 / 270

第十章 燃料——数据 / 277

机器人的自我管理 / 282

预测交通情况 / 285

个人隐私 / 290

道德伦理规范 / 291

第十一章 涟漪效应 / 297

工作岗位 / 303

交通事故带来的工作收入 / 307

新型车身 / 310

新的市场 / 312

新型的零售业 / 313

犯罪与（成人）娱乐 / 317

未来之路 / 319

后记 寒武纪大爆发 / 323

机器人的进化 / 326

传感器王后 / 328

幂次增长之王——算法 / 332

算法的进步大量涌现 / 334

致谢 / 337

前言

普通汽车终将退出历史舞台。

得益于移动机器人技术的快速发展，汽车即将成为我们可以放心托付自己性命的第一代自主式机器人。在经历了数十年不断失败的尝试后，借助速度更快的电脑、可靠的硬件传感器，以及被称为“深度学习”的新一代人工智能软件，汽车可以获得与人类相似的能力，在无法预测的环境中自主安全驾驶。

本书就是为了讲述这场变革。我们之所以对无人驾驶如此关注，原因有两点：

首先，具有巨大影响力的新技术一直是我们的关注重点，而无人驾驶汽车很有可能成为我们在有生之年所接触到的最具颠覆性的新机器之一。

第二个原因则更加偏向个人需求。就像大多数人一样，无论天气如何，我们每天都要开上一两个小时汽车，车上经常还载有“贵重物品”——孩子、朋友，或者宠物。但其实若有可能，我们并不想坐在方向盘后面，而

是更愿意单纯地享受汽车为我们提供的私密性空间和便利性行程。毫无疑问，当谷歌的无人驾驶汽车在几年之前展现出确凿的技术进步之时，我们就开始紧密关注这个领域了。

未来的几十年里，无人驾驶汽车将逐渐取代由人类驾驶的汽车。当运输行业演化成为一种自动化、即时需求的服务时，世界各地的汽车运送人类和货物的方式就会产生翻天覆地的改变。无人驾驶汽车将会转变我们对时间和空间的认知，如何出行去上班、住在哪里、如何购物等都会受到影响。

我们相信这些转变是积极的，因为无人驾驶汽车将会拯救数百万人的生命，自动化的交通管理软件将会顺畅地引导汽车行驶并有助于治理空气污染，父母无须再每天花费几个小时将孩子载到学校或者活动地点，老人和残障人士也将能够便利出行。

每一种影响力巨大的新技术都有其不利的一面，无人驾驶汽车也不例外——无人驾驶汽车出现后，数百万的货车司机和出租车司机将会失业。公共运输行业也会凋零，因为人们都会被按需服务的无人驾驶小巴的强大便利性所吸引，它可以随时随地承载任何人到各种地方，而花费不过是一张车票钱。不过，除非有严格的隐私保护措施实施，无人驾驶汽车的乘客终将会意识到，他们其实是牺牲了自己的隐私以换取出行的安全和便利性，因为引导他们乘坐的无人驾驶汽车的软件系统会追踪并记录他们的每次出行。

在本书中，我会进一步解释汽车是如何转变成智能运输机器人的。我们会评估无人驾驶将对汽车行业产生的影响；讲述在无人驾驶改变人们的每日出行，将其由一种沉闷而危险的活动变成便及每个人且避免冲突的行

为之后，城市发生了怎样的变化；还会探究人们在自动化驾驶探索道路上近六十年的失败尝试。最终，我们会引导读者通过对硬件技术和软件技术的清晰而明确的发展性认识，了解现代的无人驾驶汽车是如何诞生的。

我们的目标，就是向读者提供他们在面对即将到来的新世界时所需的见解和认知。在未来，无人驾驶汽车的数量将会远多于由人类驾驶的汽车，我们衷心期望你会享受这趟旅程。

导语

机器司机

在不久的将来，全球汽车博物馆里展出的汽车，都不会是20世纪初之前的那些新鲜炫目的车型，就像历史的叙述者来到一个历史遗迹后，会因害怕后人精心保存下来的中世纪小教堂遗址，参观博物馆的人都会偏重这些

Introduction

感知。20世纪初的人类竟然会用地毯下看不见的那种交通方式。

我们当前的汽车是非常智能的，它有四个轮子，一个柴油引擎和一个气动发动机，这样一套标准中供自动驾驶“平台”。自100年前被引入后就从没有变过结构或性能或效率。到现在100年间，随着越来越多的智能软件，几乎无时无不在被重新设计，但其体型不断缩小，价格逐年下降的同时而在确保硬件性能强而耐用。世界上其他地区的社会都在发生着根本性的变革。

机器司机

在不远的将来，全球汽车博物馆里展览的汽车，都会变成21世纪初之前的那些光鲜靓丽的车型。就像历史的痴迷者来到一个历史遗迹后，会闪身钻入精心保存下来的中世纪小屋回廊里，参观博物馆的人也会钻进这些展览车辆的前排座椅里。这些参观者坐在方向盘后面，拨弄着内置GPS的显示器屏幕，游戏似的用他们的脚踩动着刹车，这一切恐怕会让体验者不禁感叹，21世纪初的人类竟然会用如此不便甚至危险的交通方式。

我们当前的汽车是非智能的。它有四个轮子、一个金属机身和一个气动发动机，这样一套标准化的自动系统“平台”，自100年前被引入后就从没有发生过任何本质性改进。而在这100年间，随着越来越多的智能软件，几乎无处不在的通信网络，以及体型不断缩小、价格逐年下降的强大而准确的硬件传感器兴起，世界上其他产业的根基都在发生着根本性的变革。

得益于机器人技术和人工智能软件在近期取得的成就，非智能汽车的时代终将结束。普通平凡的汽车即将发展进化成自动化的移动机器人。

在近一个世纪的时间里，由人类驾驶的汽车已经改变了我们的生活。这种无需马匹拉动的“自动”车厢的出现，已经改变了“行走之城”的城市格局。曾经如蜂巢一般，由各种小而弯曲的小巷、庭院、商铺、公共广场交织而成的城市演变成了“汽车之城”，打造出了由广阔街道和停车场组成的整齐的大网格。汽车不仅给人们带来了自由，同时也带来了新的工作机会和社交机会。汽车也给商业贸易带来了便利，商人可以快速地将产品运输到以前从未到达过的市场。

然而，如此宝贵的个人移动便捷性也让人类付出了极高的代价。在近百年的进程中，交通事故已经夺去了数百万人的生命。当汽车赋予人们开车到远方工作的自由时，也催生了一种新的问题——城市交通拥堵。今天，世界上所有城市的人们每天都要出行、上班通勤，或者通过汽车运送货物，这使得城市上方的空气层恶化成一个黄色烟雾似的“油斗篷”。

粗略估计，目前全球有十亿辆由人类驾驶的汽车在陆地上漫游。对汽车的依赖已经使我们在很多方面付出了高昂代价。但归根结底，对于世界上的大多数人而言，汽车仍是目前个人出行中最快速、最廉价、最舒适的方式。不管怎样，汽车仍将是现代生活中不可缺少的一部分。

事实上，解决汽车引发系列问题的最好方式就是让它们变得更智能。当人类驾驶者让智能软件接管方向盘的时候，无人驾驶汽车将为全世界数十亿人提供一种更安全、更简洁，甚至更方便的出行方式。在未来的十年，自动驾驶汽车将出现在全世界的大街小巷，它将再次重新编排我们的

居住环境以及工作娱乐环境。

不相信？这也难怪。

近百年来，各行各业的专家都在预测人类的霸主地位将会被智能机器所取代。但到目前为止，这些预测仅仅在一些高度细分化的工业岗位上成真了，或者是仅限于虚拟世界中的活动。例如，机械手臂可以完美无瑕地取代曾经由工厂工人实现的操作；在虚拟世界里，人工智能软件的能力已经超越了人类，它们可以玩棋类游戏，可以迅速地进行股票证券交易，或者在复杂的大运量公共客运系统中找出最优路线。

现代化的软件极大提升了人工智能的能力，先进的机器人还可以执行诸多技巧性的工作，这些都已成为现实。但是如果机器人的身体并非由螺栓插销固定，而是由仿人类的机械躯干组成，使其能在运动过程中与周边环境互动，那么即使是当前最先进的人工智能软件系统也无法顺利地承载这项复杂的管理任务。在后面的章节中，我们会陆续探讨相关的原因，但目前来看，今天那些能够移动的机器人，在身体敏捷性和感知能力上也就相当于蟑螂的水平——或者状态良好的话，可以达到蟾蜍水平。

当机器人学家还在研究移动化智能机器人的道路上步履维艰时，建造一辆踏实可靠的无人驾驶汽车从技术上而言已经触手可及，这同样也是工程学上的一项壮举。因为对于程序员为人工智能写代码操作运动过程来说，汽车相比其他形式的移动化机器人有一个巨大的优势——滚动要比行走或爬行简单得多。

如果人工机器人具备多个肢体部件，那么管理其动作快速运行的软件

从体量和复杂性上都要比智能汽车大得多，因为多肢体部件可以演化出数量近乎无限大的不同动作模式和相对位置。相反，一辆汽车的四个轮子、刹车以及方向盘组合出的运动模式几乎全部是可预测的。引导无人驾驶运动的软件所管控的动作组合数量必然相当少，例如多半只是车轮的左右转动，或者监控汽车是在减速还是加速。

驾驶行为可以自动化的第二个原因在于，开车是一个相对刻板重复的反应式活动，无论智力高低都能拿到驾照，因此无人驾驶的智能只需满足以下条件就够了：对清晰可见的路面危险——比如不断靠近的路面坑洼或者缓慢行走的成群结队小学生——做到即时反应，沿着清晰明了的道路或高速行驶，并且遵守相对简单的交通规则。

说到这儿，怀疑论者就会指出这背后一定另有故事。的确，如果只是简单地编辑程序让四个轮子的机器人遵守道路交通规则，那么无人驾驶在几十年前就早已普及了。汽车为什么直到现在才有可能变得智能化呢？源于两个原因：

第一是现实因素，也就是它的门槛很高。汽车毕竟是行驶在公共街道上的重达两吨的“铁盒子”。如果引导无人驾驶的软件出现了问题，后果将不堪设想。“人命关天”，这也就解释了为什么今天的第一批自动驾驶车辆会先在人烟稀少的地方使用，因为车辆如果在这种地方出问题、偏离既定的路线，所造成的人员伤亡是最小的。例如，在偏僻的澳大利亚北部矿石公司使用巨大的自动化驾驶卡车来运载矿石；农民使用自动驾驶的拖拉机、联合收割机在广袤而无人的农地里耕作；在配送中心和工厂里，专门的自动化车辆会将货物从房间的一侧运到另一侧；在度假村或机场，名

为纳维亚的无人驾驶穿梭车会以既定每小时15英里的速度沿着固定路线往返运送乘客……

制约无人驾驶发展的第二个难点及关键点就是纯粹的技术问题。虽然在99%的时间里，驾驶过程都是思维僵化而且可预测的，但是仍有1%的时间里会产生突发事件。生物有机体会依赖所谓的“本能（Simple）”来应对生活中突发的未知事件。是“本能”，使人类可以在交通的高峰时段自由驾驶，也正是“本能”，蕴含了人类智能中尤为难以复制的大量能力。

机器人学家对这种只在1%的时间内才发生的突发性小概率事件赋予了专门的名称——“极端事件（Corner Cases）”。极端事件并不常见，也难以预测，却可能导致灾难性后果。机器人的人工“本能”在处理极端事件时的效果最终决定了它的可靠性与价值。如果一个机器人的软件无法处理它所遇到的每一个极端事件，那么希望人们不会放任它独自行动，否则最坏的情况就是机器人无法完成交代的任务并造成了严重的破坏。

驾驶也许是最为重复刻板的活动，但同时也充满了数不尽的、潜在的致命性极端事件。管理无人驾驶的软件必须能够本能性对突发情况做出反应，例如跳上汽车引擎盖的鹿，或者向汽车挡风玻璃上喷漆的愤怒乞丐，这些都有可能发生——还要祈祷乘客会为挡风玻璃上的污渍付清洁费——所以，尽管投入了几十年的努力，自动化工程师和机器人专家还是无法编写出一个有效的软件，处理无人驾驶在道路上可能遇到的无数种突发状况。

机器人技术中一条基本的规则就是，环境越是简单、越是容易预测（极端事件越少），构建一个使机器人在这种环境下执行任务的软件就越简单。机器人能在工厂里广泛普及是因为大多数工厂都是一个封闭的世

界、一个高度结构化的环境，在那里极端事件是能够预测的，而且会被工业工程师小心地排除。在封闭的环境中，机器人的工作职能会围绕特定任务进行设计。工厂里的机器人知道将会发生什么。指导它们运行的软件会引导机器人进行一系列一成不变的活动，例如冲压金属部件、固定螺栓，或者将箱子从一侧拉到另一侧。

虽然在工厂里可以设计一个整齐划一的封闭式环境，然而在实际环境中，街道和高速公路上是复杂混乱且难以预测的。汽车方向盘后的每个司机不仅要处理新的突发情况，还必须应对其他相关挑战，这些挑战是软件程序难以处理的，由模糊的或者高度情境化的行为规则调控的信息交互问题。人工智能软件尤其会在对安全驾驶有重要影响的两类活动中出现问题，即复杂的非语言信息沟通，和在多种情境中对同一物品进行持续正确识别。

驾驶行为要求司机和行人之间进行复杂的“社交”活动。当人类司机位于方向盘后面时，通常会参与到一场非语言的“社交芭蕾舞剧”中，并用点头、挥手、眼神交流来传达他们的意图。挥手和微笑对于人类而言可能是简单的，但是想要编写一套软件来识别人类的面部表情、身体语言并做出恰当的回应，则异常困难。

无人驾驶汽车，这一移动版机器人不仅在识别复杂的非言语交流上存在困难，当需要应对突发事件时，它们的计算机智能也会捉襟见肘。这些问题均是由知觉缺陷导致的，即它们不具备对所见事物进行感知区分并合理应对的能力。理想情况下，计算机科学家会通过编写一套程序来解决这个问题，为汽车提供一致性、准确的人工意识和情境理解能力。然而直到

目前为止，这个问题还没有被解决，这样的软件还没有做出来。自从人工智能领域开辟半个多世纪以来，计算机科学家和机器学家为了实现“知觉”这一神秘技艺的自动化已经尝试了各种方法，但均未能取得突破性进展。

生物有机体的知觉系统包含了一系列广泛的能力。知觉中的一个方面就包含认知识别功能，人或动物因此具有“读取”复杂情境的能力，并知道如何正确地反应。知觉的另一个功能就是加工处理视觉信息。生命体具有高度发达的视觉系统，可以良好地识别一个物体，甚至是从不同的角度、在不同的光线条件以及不熟悉的情境中都可以准确识别同一个物体。

视觉感知需要对视觉信息进行正确识别并进行归类，人类在进行这样的过程时，纯粹依靠本能，而且几乎能够做到完美无误。然而，我们这种感知所见事物的能力似乎拒绝被自动化。几十年来，机器视觉领域的研究者一直在努力研发能够快速而准确地“理解”环境中视觉信息的软件，但均以失败告终。

在机器人发展史的大部分进程里，它们都难以处理视觉信息。工业机器人应对这一缺点的方式就是在全封闭、黑暗无光的工厂环境中纯机械化地辛勤工作。而对于那些工作中涉及某些视觉活动的机器人，它们的工作方式则被设定为绝不去分类或检查它们不熟悉的事物。

阻碍机器视觉软件发展的另一因素就是计算能力的不足。由于处理图像是一项数据密集型活动，所以第一代机器视觉系统只能采取一种结构化路径的方式来实现这一流程，即用一系列“规则”来解析视觉信息。这些早期机器视觉系统的工作原理是想尝试将识别的物体与机器人内部储存的已知物体小型样例库做匹配，但显然这是一个缓慢、不准确也不灵活的过程。

早期机器视觉软件的最大缺陷之一，就是它们在应对新鲜事物或情境时表现不佳。即便是很熟悉的物体，只要在略微不同的场景中出现，任何借助这类软件引导的机器人（或汽车）都无法进行准确识别。鉴于准确识别邻近物体这一能力对于安全驾驶至关重要，机器视觉软件的差劲表现数十年来一直阻碍着无人驾驶汽车的发展。然而，最近在人工智能领域的一项突破进展可能会彻底改变这一切。

在人工智能研究领域原地踏步多年之后，一种被称为“深度学习（Deep Learning）”的新型软件于2012年诞生了，它可以对随机的数千个数码图像信息进行准确分类，并达到人类水平的精准度。虽然对随机物体的图像进行准确分类的能力听起来无足轻重，但这一能力的确是人工知觉（Artificial Perception）的基础。一旦某个物体可以被精准识别，它的信息会被“转化”给其他类型的人工智能软件，而那些软件就可以发挥其传统优势：利用统计分析或某种逻辑、规则推导出最佳反应。

深度学习软件对于无人驾驶如此有价值，就是因为它可以在开阔公路这种非结构化的环境中充分发挥作用。深度学习隶属于“机器学习（Machine Learning）”这一类人工智能软件。这类软件的突破创新并非由人类程序员来设计，机器学习软件并不会去建立一个现实世界的模型，再用正常的逻辑和规则来处理其中的问题，相反机器是通过大量的训练数据后被“喂养”出来的。例如，为了开发无人驾驶所用的深度学习软件，程序员每天会用几千兆字节（Gigabyte，即GB）的原始视觉素材来“喂养”这个软件，而这些视觉素材都是通过车载摄像头在真实道路环境下的行驶中拍摄收集的。

深度学习软件通过观察世界来“学习”，这一点赋予了它另一巨大优势——不依赖规则（not rule-bound）。人类的婴儿学会依照物体的突出可识别性特征来识别物体，同理，深度学习软件在对物体分类时也是依据物体的视觉特征。使用规则依赖的传统途径编制的软件，遇到坐在自行车上的猫的图片时，就会陷入困惑。相反，深度学习软件会关注猫身上可供识别的视觉特征——尖耳朵和尾巴，然后不管这只猫是出现在一个多么奇怪的情境中，软件都会迅速而准确地推测出它只是一只猫。

深度学习软件已经彻底改进了对人工知觉的研究，并极其成功地应用到语言识别领域，以及其他需要软件来处理古怪异常信息的活动中。过去的几年里，为了寻求深度学习的专业技术支持，很多汽车公司都在硅谷设立了分部。谷歌和百度这些软件巨头本身已经具备了管理巨量信息和构建智能软件的能力，再加上深度学习领域的专业优势，足以支撑它们与曾经不可一世的汽车巨头在无人驾驶领域展开角逐。

深度学习对于人工智能领域的影响极其深远，甚至在我们写作本书时，其影响还在逐步拓展，而且未来几年可能会持续扩大。深度学习所引领的技术领域并不只是汽车一隅。我们预计，深度学习将会对移动类机器人的总体发展轨迹产生变革性影响。当机器人获得了通过视觉理解周遭环境的能力时，人工生命形态的发展可能会开始延续5亿多年前有机体生命发展的路径——化石显示大约在寒武纪以前，所有形态的有机体生命都是近乎眼盲的。当5亿多年前寒武纪开启时，那些近乎眼盲的有机体突然神秘地演化出了复杂的新型视觉系统，要知道先前它们的“眼睛”只是些光感性的原始细胞集。一旦它们具备了看的能力，这些简单的有机体随之演化出了复杂的身体形态，以便实现快速的反应和移动。反过来，新的身体机能

又会需要发展一个更大号的大脑来监管这些新生肢体的协调。当装备上视觉系统、快速响应的身体以及更大的大脑时，曾经不起眼的细胞集演化出了种类繁杂的复杂生物，它们从最初生命的泥沼中爬出，并努力在陆地上寻求属于自己的一块领地。

关于寒武纪生命大爆发——发生在寒武纪时期的生命大量快速演化，有一个有趣假设，即“光开关理论（Light Switch Theory）”。这个理论是由安德鲁·帕克（Andrew Parker）提出的，该理论认为正是眼睛的进化，引发了生命体之间的进化竞赛，使得那些具备最佳视觉的生物才最有可能生存下来。也许“光开关理论”理论对于机器人也是适用的。

曾经“眼盲”的机器获得了知觉的能力，它们也会从原始的禁锢中爬出，摆脱现今我们为其设计的结构化的、黑暗的工厂环境。健全的机器视觉将使得机器人可以充分利用其新添加的身体部件——轮子、四肢或踏板，而那些部件能为它们提供更高水平的灵敏度。想要控制复杂的机械新“四肢”，它们的机器大脑也会随之扩张。当我们目睹机器人掌握了新的技能并找到可以发挥功效的新应用场景时，我们也终将见证机器人形态和功能的“寒武纪大爆发”。

迎接无人驾驶

一群热带鱼的行动过程十分值得观赏。它们组成紧凑而均匀的队形，扭动着明亮鲜艳的身体在水中游弋。这几十条不同的鱼可以一瞬间同时转

向掉头，动作一致犹如一个整体。如果有个障碍物突然出现在鱼群的行进路线上，鱼群会分开绕过，然后迅速重新建立起它们先前的队形。鱼群彼此之间从不相撞，也不会触碰到浪涛裹卷而来堵在路上的异物——枝条、石子或珊瑚礁。

在理想的将来，我们的街道和高速公路上会充满成群的、紧密分布的无人驾驶汽车。与鱼群一样，无人驾驶汽车的车流会展现出惊人的防冲撞能力，在充满行人的城市街道上机智而快速地穿梭，在漫长而空旷的高速公路上以经济油耗的形式灵活停靠。有些车会携带一两名乘客，还有些车是完全没有乘客的，那是因为它们要去送外卖或者去幼儿园接小朋友。

那么，从今天人类掌控的这种糟糕交通状况，到实现理想的未来——各种形状、大小的无人驾驶汽车井然有序而安全地填充在道路上，我们应该如何着手呢？在欧内斯特·海明威（Ernest Hemingway）1926年的小说《太阳照常升起》（*The Sun Also Rises*）中，主人公比尔问麦克：“你是怎么破产的？”麦克回答：“两种方式，慢慢地，且突然地。”

实际上，技术的发展是迅速的。目前为止，读过本书的绝大多数人还是熟悉摩尔定律的。摩尔定律指出当计算机芯片的表现以幂次方提升增长时，芯片的价格和尺寸也会以幂次方的速度大幅降低。正如摩尔定律反映的，构成无人驾驶汽车的传感器、海量数据以及处理和计算所有信息的处理芯片，这些技术已经变得相当成熟可靠、物美价廉了。无人驾驶汽车的具体配置可能千差万别，但大多数无人驾驶汽车都会装载几个数码镜头、一个雷达传感器以及一个激光雷达装置来“看看”它们现在在哪里。数字化的汽车在配置全球定位系统（GPS）的同时还会匹配另一个称为惯性测量

单元（inertial measurement unit, IMU）的定位装置，用以弥补GPS的误差。还会有一个车载电脑负责将各个传感器和GPS的信息汇总，将数据传递到一个包含各个交叉路口和红绿灯信息的高精度数字地图上，然后将所有这些信
息生成一个汽车外世界的数字模型，被称为占据栅格（Occupancy Grid）。

无人驾驶技术几近成熟。埃隆·马斯克（Elon Musk）是特斯拉汽车公司的CEO，十分推崇全自动化的汽车，他对当前形式的总结是：“问题比人们想象的要简单……但也不是一个人苦干三个月就能解决的，它更有可能是需要数千人共同奋斗两年才能实现。”虽然这种技术几近准备就绪，但是这一独特技术所依存的社会环境可能还未准备妥当。

诸多社会因素制约了无人驾驶的推广应用。软件开发人员不得不面对的一个问题就是用户本身。当新的软件引入到一个组织中时，影响软件普及的最大阻碍通常不是软件自身的工作表现，而是人们对之前产品的依赖。由于这些用户的组织文化和工作流程都建立在之前的软件产品上，而改变人们的工作习惯就会引发阻抗。一旦工作流程改变，有些人就会失去地盘，有些人会被迫重新思考如何做事，诸如此类。人员问题通常是隐藏在海平面下的冰山，它会阻挠一个组织成功吸收新技术，尽管这一技术可以节省组织的时间、金钱，并提高生产力。

在对无人驾驶汽车的接纳上，人员问题可能有来自消费者一方的对抗，但我们预计事实并不尽然。虽然汽车公司的高管们曾大胆地坚持认为人们喜爱驾乘的体验，而且会继续偏爱自己驾驶汽车，但是我们认为消费者的接受度并不能成为一种阻碍。

毕马威会计事务所（KPMG）的调研指出，当技术成熟且没有人身安全

的风险时，更多消费者乐于乘坐无人驾驶汽车。当调研人员要求人们从1到10评估他们使用自动驾驶汽车来日常出行的意愿时，访谈小组成员的平均意愿程度是6分；而且如果无人驾驶汽车能减少人们一半的出行时间，并在预定时间内将人们送到目的地，那么受访者的意愿将会提高到8分。波士顿咨询公司（Boston Consulting Group）的调研结果也反映出了人们对无人驾驶技术的热情。在针对1500多名美国司机的调研后，波士顿咨询公司发现：有55%的受访者“愿意”或者“十分愿意”在5年内购买一辆半自动化的智能汽车；有44%的受访者表示只要经济允许，他们“十分愿意”在十年内购置一辆全自动化驾驶的汽车。报告中预计，第一代可家用的自动化汽车将会在2025年左右上市。到2035年时，市场上销售的汽车中将有10%是完全自动化的，占据价值380亿美元的全球市场。

平均来看，消费者越年轻，他对于无人驾驶汽车的热情越大。哈里斯民意调查（Harris Poll）项目对四个年龄段人群使用无人驾驶汽车的态度进行了调查，他们分别是千禧一代（18~37岁）、被遗忘的一代（38~49岁）、婴儿潮一代（50~68岁）以及老去的一代（69岁以上）。老去的一代人中有一半以上的人强烈反对：“我绝不会购买或租用自动驾驶汽车。”相比之下，千禧一代中只有20%的人做此反应。千禧一代中有25%的人表示只要确信“漏洞”已经被解决且价格合理，他们就会购置一辆无人驾驶汽车。

所以，我们认为只要证明无人驾驶汽车的驾驶安全性比人类驾驶更高，人们就会比现在更容易接纳无人驾驶汽车。

年轻一代的驾乘人员并不像他们祖辈那样迷恋驾驶，他们很乐意让机器人来开车。我们曾参加过2014年的无人驾驶汽车大会，会上来自咨询公

司JD Powers的一名高管分享了他们公司的调研结果，即人们对于汽车和驾驶的态度存在代际变化：30岁以下的人群，在人口学领域被称为“Y世代”，他们认为开车属于浪费时间，是需要他们被迫从社交媒体和互联网的享受中挤出时间来做的事情。这位主管对这一情形做了总结：“生于Y世代的这群年轻人对于这种想法——能够驾驶汽车是我们应该珍惜享受的一种生活方式——已经越来越不以为然。他们的主要目的是到达想要去的地方，他们更想让自己的时间以自己喜欢的方式发挥作用。”

事实上，制约无人驾驶广泛普及的最主要人为因素是政府的监督与管理措施，对于美国而言，就是州政府和联邦政府的交通法规、责任法以及保险承保范围。目前来看，无人驾驶发展道路上最大的推动力来自企业，而美国联邦政府对无人驾驶的监管和限制条例也开始缓慢地实施和制定。尽管如此，2016年美国交通运输部（USDOT）还是对无人驾驶的应用前景释放了积极信号，宣布对各个州的机动车管理局在无人驾驶管理方面提供指导建议。截至本书开始写作之时，美国已经有四个州——加利福尼亚州、内华达州、佛罗里达州和密歇根州——会颁发官方的无人驾驶许可证，同时也有几个州在考虑效仿。

无人驾驶许可证是一个良好的开端，但是监管应以何种方式进行还需要进行大量的研究和探索。理想情况下，政府的最高层应该制定前瞻性策略，而不是被动响应。例如，法律专家应该重新检视责任法，以明确当无人驾驶造成事故时，谁应该承担什么样的责任。可能的话，甚至有必要修订相应法规。汽车保险业也将需要进行类似的重新定义；立法者需要明确一辆汽车要达到何种程度的“安全”，才能在无人驾驶的情况下上路，以及如何检验这种安全性。监管上的这些疑虑最终可以完全解决，但鉴于它

们目前还存在着，由人类驾驶的汽车还将继续发挥余热，同时人类要咽下苦果——承担生命的失去、时间的浪费以及油料燃烧对环境的破坏。

在无人驾驶技术成熟和人为制约因素初见端倪的这段时间里，延误这项技术发展的代价几乎可以直接以人命数量计算。根据世界卫生组织的报告，车祸是导致全世界15~29岁的青年人早逝的最高因素，同时也是造成各年龄组人群死亡的第二关键因素。这些车祸中的大部分并不是汽车某部件失灵造成的，而是由本可杜绝的人为因素导致的，或者称为“4D”因素——醉酒、吸毒、瞌睡以及走神（Drunk, Drugged, Drowsy, Distracted）。只要方向盘还掌握在人类手中，车祸造成的死亡率就可能持续上升。新兴经济体中的人群刚刚尝到拥有私家车的乐趣——在中国、印度、俄罗斯以及巴西这样的新崛起国家里，随着更多汽车上路，车祸造成伤害或谋杀的人员数量将会继续攀升。

此外，人类驾驶汽车还有一个风险因素——分心驾驶也在持续增长，在2013年，因为分心驾驶，全美国有超过42.4万人在车祸中受伤，这一数据相比2011年增长了近10%。

致命病症的特效药

关于汽车，有一个奇妙的讽刺，虽然汽车自从发明以来已经“谋杀了”数百万人的生命，但我们的社会对于它造成的死亡人数仍视而不见，这或许是一种冷漠的默许。每年，全世界都有将近120万人死于车祸，这个

死亡率相当于每年释放10个广岛级别的原子弹爆炸。

与战争、暴力、毒品一样，汽车对人而言也有着致命危害。谋杀、自杀以及战争每年预计会造成160万人死亡；由毒品导致的死亡人数每年约有18.3万人。尽管全球死亡人数有很大一部分来源于本可避免的汽车事故，但是仍没有哪个政府支持“向汽车宣战”之类的活动，或者说呼吁人们停止开车。尽管高速公路上一再发生连环相撞事故导致数十甚至上百人进了医院，但是公众仍然没有对汽车公司发起持续的质疑和谴责。

如果有一种方法能减少每年死于车祸的人数，又会怎样？如果这种方法存在，那么联邦政府、州政府以及市政府就应该联合起来全面推行这个方法，减少车祸死亡人数将成为政府的头等大事。对这一方案的推广将会吸引大量资金募集，并对做出显著成果的单位或个人颁发奖章以大力表扬。类似于对乳腺癌研究作出贡献的人会颁发粉丝带徽章，联邦政府也要慷慨地为大学提供科研基金。

事实上，这种方法是存在的。它需要的就是将人类驾驶员从方向盘后面移开，并用智能软件和传感器来替代。如果我们的社会能齐心协力地将无人驾驶发展成为类似于“阿波罗计划”的下一个文明里程碑，那么我们将拯救数百万人的生命。伊诺运输中心（Eno Center for Transportation）的研究曾作出预测，如果路面上的汽车能有90%变成自动化驾驶的，单单在美国，就可以将一年里死于车祸的人数从32,400人减少到11,300人。

然而讲到此处，我们必须负责任地指出，用机器人替换人类驾驶员并不会一劳永逸地解决所有问题。

人类可能会轻而易举地找出新的方式来伤害自身。有些分析师就指出，自动化汽车提供更多安全性的同时，也可能招来新的风险，例如黑客入侵、网络病毒。还有恰恰因为乘客觉得安全，他们可能会做出新的冒险行为，例如不系安全带，或者纯粹为了好玩而在成队列行进的自动驾驶汽车之间来回穿梭。

但即使将可能出现的新型行为问题纳入考虑中，无人驾驶汽车依旧会使路面更安全。而且，安全因素并不是无人驾驶所带来的全部收益。如果人们能获得安全便利的私人交通工具，那么他们在选择到何处定居、工作和游玩的机会也会增多。无人驾驶会使出行变得更加高效，这样一来，交通拥堵和伴随的空气污染问题将会相应地得以缓解。不用再束缚于沉闷的上下班时间，可以说是无人驾驶应用后最直观的益处之一。另一个益处就是有更多的人可以享受便利的个人出行。

据美国运输部数据显示，每天都有586个老年司机在汽车事故中受伤。不幸的是，由于生活质量与个人出行紧密挂钩，所以禁止驾驶的决策常常会招致人们的抵制。实际上，老年人、视力受损的人或因其他身体原因不能亲自驾驶汽车的人，都可以通过无人驾驶技术获得按自己意愿出行的能力。

阻碍无人驾驶技术发展的7个误区

如果不积极响应无人驾驶技术的发展，我们会继续付出生命和时间的代价，继续承担污染的危害，继续丧失机遇。然而，并不是每个人都认

识到了无人驾驶汽车的价值。在写作本书时，我们发现有些针对无人驾驶的误解还在广泛肆意传播，并且这些信息会被反对者拿来对抗无人驾驶的推广政策。我们将这些误解谣传总结成7个误区。具体如下：

1. 自动化驾驶技术会脱胎于当今的驾驶员辅助技术。

有些人认为，汽车的自动化转变会体现在驾驶过程中分阶段进行，而这些阶段只是逐步拓展驾驶员辅助技术的应用范围，例如自适应巡航控制（adaptive cruise control, ACC）和车道保持辅助系统（lane keeping assist system, LKAS）。实际上，这种分阶段操控不仅存在技术难题，而且还是不安全的。研究发现，当人类和机器共享方向盘的操控权时，如果要求人类在紧急状态下突然掌握方向盘，他们反而会无法胜任，表现为注意力无法集中。另外，完全自动化驾驶所需的技术并不是当今驾驶员辅助技术的派生，而是有着独立的技术发展路径。

2. 技术的进步是线性推进的。

在预测无人驾驶的发展态势时，有些人认为机器人技术在将来的发展速度会继续延续过去十年的步调。可是，无人驾驶技术的发展只会更加迅速，因为它们所利用的使能技术（Enabling Technology）会遵循摩尔定律，即硬件性能提升的同时价格大幅下降。结果就是，随着无人驾驶汽车各组成部件价格的降低，它们的性能会迅猛提升。加速无人驾驶发展的另一力量则来源于“车队学习（fleet learning）”。由于无人驾驶汽车会将它们的驾驶“经验”数据进行汇总共享，所以每辆汽车都能从其他车辆提供的共享经验中获益。用不了几年，引导无人驾驶的操作系统就会积累起丰富的驾驶经验，其丰富程度不亚于数千个人类驾驶员的综合。

3. 公众会抵制无人驾驶。

汽车巨头们的广告部门都乐于强调人们喜爱驾驶的体验。事实上，大多数人的驾驶体验中都会涉及每天通勤上班，或外出办事时烦闷的开车时光，即在拥挤的道路上一步步挪动。绝大多数人都十分愿意将花在开车上的时间用来做点别的事。一旦技术成熟，无人驾驶汽车可以被安全使用时，消费者将会无比渴望得到它们。

4. 无人驾驶的发展需要在基础设施上进行大量投资。

实际上，引导无人驾驶软件的只是清晰的道路标志线，除此以外，无人驾驶并不需要什么特殊的基础设施建设。导致这一误解的原因已经有数十年的历史，早先美国运输部集中了大量资源来提升互联网汽车（Connected Car）的视觉识别能力，所用的方式就是在汽车和路边建筑上加装昂贵的无线信号发射器来分享数据。这种视觉识别系统不仅是无效的，还是造价高昂的。相反，无人驾驶汽车会利用机器人技术中的机器视觉，将数字地图储存到汽车中，而不是路面上，这样就将智能赋予了汽车。事实上，无人驾驶需要的基础建设投资很少，因为它们连路标和信号灯都不需要。

5. 无人驾驶会造成道德困境。

实际上，无人驾驶在道德判断方面与人类并无二致。而反对者之所以宣称无人驾驶造成了道德难题，只是源于一个令人不舒服的事实——引导汽车的智能软件在必要时会对人类和动物的生命价值进行量化比较，以便于无人驾驶汽车在应对将要发生的事故时做出恰当的反应。无论出于什么

原因，在车祸发生前的一刹那，我们人类都还是偏向于由人类而非机器来做出快速的、非正式的成本收益分析。可是，在面对这种驾驶过程中的瞬间决策时，人类驾驶员早已本能地估量出“谁可以牺牲掉”了。保险公司也早已对我们身体的各个部分出现问题时所需支付的潜在费用做出了量化。

6. 无人驾驶汽车需要有近乎完美的驾驶记录才能确保足够安全。

其实，只要无人驾驶汽车的安全驾驶记录超出人类驾驶员的平均水平，它们就能造福人类。平均来看，每20万英里，人类驾驶员就会出现一次非致命性碰撞。既然很多人可能对于无人驾驶汽车的这种“异型”智能持有偏见，也有很多人认为自己的驾驶技术在平均水平之上，那么我们认为，只要无人驾驶技术在安全性上优于人类驾驶员两倍，就该让它自行驾驶，乘客可以忙些别的事情或者睡觉就好了。这样就可以说安全系数达到了2.0，或者说“每发生一次碰撞的平均距离”达到了40万英里。

7. 无人驾驶汽车的应用将会突然大量爆发。

有些人喜欢问到底“哪年”无人驾驶汽车会完全接管道路。其实，无人驾驶汽车的应用是逐步发展的。自动化的汽车会首先应用于一些限定的区域，比如度假区、大学校园，以及封闭性的市中心区域。在公共道路上，无人驾驶的“入侵”是逐步化的，随着由人类驾驶的汽车逐渐退役，无人驾驶汽车的应用比率每年也只会增长几个百分点。

无人驾驶技术发展的时间表

无人驾驶汽车的推广并没有一个简单明确的时间表。实际上，向无人驾驶汽车世界的转变将是逐步推进的。汽车并不会在某个特定的年份一下子都变成无人驾驶。有两个原因：首先，无人驾驶汽车的应用会先发生在某些场所或某些国家；再者，汽车公司目前正在采取的策略分阶段自主化。如果他们成功了，人们可能需要在行驶过程中某段时间来驾驶，这也使得汽车向无人化的转变不可能存在一个明确的时间节点。

第一代自主化驾驶汽车在行驶上普通街道前，会先应用到某些特殊场景中。矿场和农场已经开始使用自动化汽车了；货运卡车也将是早期的应用领域之一。

最初，人们对无人驾驶汽车的应用非常谨慎，主要是以低速行驶的穿梭车形式存在，应用于封闭的、固定化的环境中，如机场和度假村里。英国的米尔顿·凯恩斯镇（Milton Keynes）正在测试自动化的双座电动出租车，它将会在人行道和乡间小道上运输乘客。随着时间的流逝，如果这些自动化汽车表现良好，并且安全性得以验证之后，它们的运行速度和行驶区域将会逐步扩展。我们预计谷歌公司销售的第一代无人驾驶汽车并不会成为消费者的日常代步工具，反而更有可能成为企业或小城市的补充性运输方案。在某一时刻终将会迎来平稳的过渡，自动化穿梭车会驶出封闭的区域，踏上城市公路。

向无人驾驶汽车过渡过程中的另一个方面就是区域位置，即究竟哪个地区会率先将无人驾驶汽车应用到日常生活中。有些国家会早于其他国家推广无人驾驶汽车。在同一个国家里，有些州或省份又会早于其他地区先对无人驾驶汽车进行立法管理。

想要明确无人驾驶汽车的“完全成熟”时期，那就要明确什么是“完全的自动化”。完全的自动化是指道路上100%的汽车都是全自动化的，并且是全天候的。以这种方式来定义完全自动化就意味着这个过程会持续至少一个世纪之久。因为无人驾驶的应用领域很多，人们对孰先孰后并没有形成一致观点。

应用时间表中的诸多变数，有一部分源于现实的考虑。因为汽车必须满足严格的安全标准和排放标准，所以汽车领域的新技术在普及应用时要比其他领域的技术缓慢得多。另外，汽车价值不菲，因此人们对汽车的留恋会持续数年。人们添置或舍弃汽车的速度要比更换智能手机的速度慢很多，这也就导致从人类驾驶向无人驾驶的转变将会跨越数十年。

总之，汽车公司和运输部官员也会用更长远的眼光看待这件事，他们预计无人驾驶汽车想要成为公路上的主流至少也要在2025年之后。根据汽车市场研究公司IHS的预测，第一代面向市场销售的自动化汽车将在2025年前后出现。IHS公司估计到2035年时，销售的新车中将有大约10%是自动驾驶的，每年产量总数可达到1180万辆。2050年之后，售卖的所有新车将几乎都是自动驾驶的。

通常，汽车公司也偏爱以分阶段的方式逐步推广自动化驾驶，这也是我们无法准确界定无人驾驶汽车全面应用日期的另一个影响因素。汽车公

司为了推广驾驶员辅助技术，会在新闻稿的标题上吹嘘“某某汽车公司将在2020年发布无人驾驶汽车新产品”。然而，仔细阅读后会发现这个公司所宣传的产品，实际上只是具备了一些使汽车可以在某项特定的任务中自行完成的功能，例如在某些条件下自动泊车，或者只是对巡航定速控制和车道保持技术结合后产物的过分美化。

对于汽车能够在各种环境下全自动化驾驶的时间表，科技公司会作出相对乐观的预计。谷歌和特斯拉公司都十分坚定地相信未来的驾驶工具一定是全自动化的，尽管确切的时间和细节还有待确认。2014年10月，特斯拉公司的埃隆·马斯克在接受彭博电视台（Bloomberg TV）采访时就讲道：

“从现在起五六年內，我们就能造出真正的自动化驾驶汽车，那时你就可以直接上车睡觉，醒来时就到目的地了。”但他也强调：“想要得到监管部门的审批认可，可能还需要再花上两三年。”

分析师托德·利特曼（Todd Litman）预言，如果没有联邦政府的推动应用，无人驾驶汽车的普及将会延续自动挡变速器的推广进程，那个过程则要耗费将近50年。利特曼估计即使无人驾驶汽车在2020年时就合法化了，它的普及应用也还是需要几十年的时间。他预计到2050年时，无人驾驶汽车将占据新车销售数量中的80%~100%。可即便如此，路面上行驶的汽车中仍有40%~60%是由人类驾驶的。

即便是无人驾驶汽车明天就面市了，将世界上所有人类驾驶的汽车淘汰也不是一件小事。在美国，有大约2.5亿辆汽车在路面上跑着，因此很多分析师将美国称为“活动的停车场”。每年，这个活动的停车场中有1300万到1400万辆的汽车会被淘汰掉，扔到垃圾场里。即使人们可以直接购买

一辆经过充分测试并合法认证的自动化汽车，但由于现在汽车的平均寿命是10~15年，所以想要将所有老式的、人类驾驶的汽车撤离道路仍需大约20年的时间。

无论谁的预言成真，有一件事是始终确定的，那就是向完全自动化驾驶汽车时代的演变将历时几十年。尽管谁负责驾驶什么样的工具，以及在何时何地行驶等细节还有待明确，但可以确定的是人类和机器人在未来的几十年里会共同掌控道路。

在后面的章节里，我们会从多个角度探索无人驾驶汽车，揭露那些挡在无人驾驶汽车发展道路上的谣言。我们会重新打造城市格局，将停车场改造为人们休闲的空间，通勤也不再那么痛苦。对于赋予无人驾驶汽车“看、反应、思考”能力的机器人技术，我们会深度剖析。我们会探究汽车、媒体以及零售业将因此迎来怎样的变化。今天的自动化驾驶汽车是人工智能和机器学习领域研究者们数十年科研的成果汇聚，对于前人这段漫长而丰富的研究过程我们也会一起重新领略。

第一章

无人驾驶的世界

A Driverless World



如果将世界各地街道上行驶的数十亿辆汽车都奇迹般地转换成可靠的无人驾驶汽车，你首先会注意到的事情是世界变得安静了。毕竟，咆哮的警笛声和刺耳的喇叭声只有在方向盘前坐着人类驾驶员时才能发挥作用。那时街道上遍布的，也多是类似高尔夫球场上代步车的小型车，有的搭载一两名乘客，有的完全是没人的。时不时有类似房车的车辆堂而皇之地驶过，那可能是供上班族使用的设施齐全的移动办公室。偶尔，也会有一两辆人类驾驶员开着的车出现。当人类驾驶的车辆出现时，其他车辆会收到提醒，变得谨慎起来，与人类驾驶车辆保持很宽的安全距离。

想要叫出租车，你只需在手机上按一下。几分钟后，一辆无人驾驶移动座舱就停到你的身边。如果你同意共享行程，那么你呼叫的座舱里可能已经有一两名跟你去往相同方向的乘客，这种轻微的行程打扰却能大大减少你的乘坐费用。

就像电梯的内部一样，这个移动座舱里光秃秃的，一切简化，只有坚

硬并易于打扫的陈设，几乎没什么能移动的部件。与高峰时段过后的破旧地铁列车类似，无人出租车的地面上也常常散落着食品袋子和各种烟头。出租车的座椅上还粘着嚼过的口香糖残渣，覆盖着肆意的涂鸦乱画。出租车里的安防摄像头完全看不到这些乱丢的废物和受到破坏的车内环境，因为它的镜头早就被一个廉价的针织帽盖住了，也就一直处于眼盲状态。



图1.1 个人定制化的移动办公室座舱（概念图）

来源：IDEO设计公司

在这个新时代里，乘客给人类司机付现金的行为早已不复存在，那种不具名的出租出行也已成为过去技术低端的代表。现代的移动座舱在你进入的那一刻，就知道你是谁了。既然你允许身份识别，移动座舱就会迅速检查你的网页浏览记录、购物历史以及最近的行踪，并向你提示在去往目的地的路上会经过两个你喜欢的商店。

除去车内不停播放的广告娱乐信息有些烦人外，整个行程还算是放松

的，无需什么人际交流。正如在地铁里一样，乘客之间会尽量避免眼神交流。因而，也就无须再像从前那样与人类司机闲谈。既然你相信车载软件会规划出最便捷的路径送你到目的地，你也无须偷偷地在手机上查看地图比较路线。

到达目的地时，车费会自动从你账户上扣除。没有司机，也就不用额外支付小费。还因为你与其他人共享乘车，承受广告的无尽骚扰，车费更是便宜。鉴于出租车内陈设简单，虽然车内并不是很干净，行程也不是很舒服，但整个过程却很简便。就像是乘坐电梯一样，乘坐出租车出行不再有各种冲突麻烦。

便利的个人移动性

无人驾驶汽车对人类生活产生的影响中最不为人全面认识的一点，就是对于交通拥堵产生的影响，以及随之而来的弊端。一种乐观的局面是无人驾驶汽车可以提升城市交通系统的效率，致使私家车数量减少，进而减少拥堵，降低城市中交通出行产生的碳排放量。另一种对环境不利的局面是，在人们享受着个人移动性的便利的时候，平均而言，无人驾驶汽车实际上每年将会刷新出更高的车辆行驶里程，也就导致了更多的碳排放。

出行的便利是把双刃剑。人们就像铁屑被磁铁吸引一样，为出行的便利性所吸引。然而，随之而来的还有一些代价以及未曾预料到的负面效应。无人驾驶提供的个人出行便利性也许能解决现代汽车技术加于我们身上的资源严重过剩问题。另一方面，它可能会带来隐藏成本——越来越多

的人会在不经意间突破他们原本的出行行驶里程。这种新技术应用所带来的期望效益不增反减，是源于效率提升反而促进了使用率，经济学家将这种现象称为“回弹效应（Rebound Effect）”。

无人驾驶汽车的出现是否会对交通状况产生“回弹效应”，即增加了人们每年乘车出行的里程数，并导致道路上的汽车数量不减反增，这一点还尚未得以明确。不过有些研究给出了一个乐观的场景，那就是在未来的几十年里，城市街道上的车会越来越少。国际交通论坛（International Transport Forum）是专门提供交通政策建议的一个智囊团。智囊团成员路易斯·马丁内斯（Luis Martinez）在接受《经济学人》（*The Economist*）杂志采访时预测，将来自动驾驶车辆组成的车队会取代城市中所有以车辆为载体的公共交通，如出租车和公交车，我们可以用更少的车辆为人们提供更多的出行便利。

为了验证这个理论，国际交通论坛的研究者创建了一个以个体为基础的模型，用来模拟欧洲一个中等大小城市的日常出行模式。通过将先前交通状况调查所积累的数年真实数据导入模型，研究者计算出一个结果——如果城市居民不再乘坐私家车和公共交通，都是以共享的方式使用自动化出租车出行，那么城市道路上的机动车数量会减少90%；虽然成群出动的自动化汽车会极大减少路面上汽车的数量，但是模型计算后还发现每辆车的出行总里程会略微增加，这是因为自动驾驶的出租车会更频繁地穿梭于道路上以接送不同乘客。

密歇根大学的交通研究所（University of Michigan Transportation Research Institute）给出的一份报告就支持了上述模型的结果。报告指出，自动驾驶汽车可以减少美国家庭平均占有的车辆数量，从每个家庭两辆以上减少到

只需一辆车。根据报告所言，一个家庭拥有一辆车之所以成为可能，是因为自动驾驶汽车在送一个家庭成员去上班后会启动“回家”模式，这样其他家庭成员也可以使用这辆车来接送自己外出办事或参加活动。

然而，还是存在一个问题。虽然一台家用无人驾驶汽车可以来来回回高效率运送家庭成员，但事实是一辆车服务的人员越多，它的行驶里程也就越多。尽管未来每个家庭平均占有的车辆更少了，可是留下来的无人驾驶汽车在使用频率上要增加75%，累加起来就是每年每辆车平均里程达到了20,406英里。这一事实的积极面在于，即使一辆无人驾驶汽车平均而言要多跑75%的里程，但整个家庭的总行驶里程还是要比家中有两辆人类驾驶车辆使用时的里程更少。但这个更有效的交通出行会引发未曾预料的负面结果，即：一辆无人驾驶汽车的行驶里程数要远远高于一辆人类驾驶的汽车。召唤一辆无人驾驶汽车来接送你，无疑会带来极大的便利。然而，这也可能意味着召唤来的汽车要大费周折地多跑两倍里程来接你。

理想状态下，一辆自动驾驶汽车在空车状态时会驶离交通干道找一个安全的地方停泊，等待下一次的召唤。可是，如果这个安全区离街道有几英里远，那么相比就近停靠而言，无人驾驶汽车就要被迫空载来回行驶很长一段距离。它的行驶里程会增加，而那些浪费的折返行程加剧交通拥堵和空气污染状况。

如果人们在获得便利的个人移动性的同时会导致每年个人出行里程的大幅增加，那么无人驾驶汽车的出行就会对环境产生破坏性影响。今天，交通出行已经成为空气污染的最主要来源之一。单单是在美国，汽车和卡车排放的尾气就占到了每年人类活动产生温室气体排放量的大约29%。如果

无人驾驶汽车会导致人均车辆里程数的增加，那么发展中国家里人口稠密的“特大城市”所面临的问题会更加严重。

尽管美国的汽车发展已经有将近100年历史，其他的高速发展国家也在积极地迎头追赶。中国就在追赶美国的脚步，逐渐建立自己的汽车文化。随着越来越多的中产阶级新贵享受汽车出行的便利，类似北京、郑州等城市不得不忍受那惊人的八车道交通拥堵和日益严重的雾霾。

时至今日，中国的人均车辆占有率还是要比美国或欧洲低不少。在中国，平均每1000人拥有85辆汽车，而美国则是797辆。然而，中国汽车行业每年制造和销售的新车比率还在不断刷新纪录，自2013年以来已经达到7%的年均增长率。通过尽早应用无人驾驶技术，也许中国的汽车文化能规避某些资源的严重过剩问题。为了改善交通状况，被誉为“中国谷歌”的搜索引擎公司百度已经发表声明，称会与宝马公司联合开发专门适合于中国道路的自动驾驶汽车。

无论是在发展中国家还是发达国家，交通拥堵都是空气污染的主要来源。在美国，由于上班族在交通拥堵中寸步难行，他们汽车的发动机空转每年会浪费29亿加仑的汽油，覆盖的面积足以填满4个足球场。无人驾驶汽车可能会行驶得更有效率，产生更少的污染，但也可能促使人们每年行驶更多里程，使得空气质量更加恶化，城市交通拥堵更加严重。

汽车每年行驶里程总数的增加对环境产生的另一方面影响可能来源于无人驾驶汽车有着更短的使用寿命。一辆汽车的寿命可以由其里程表预测，根据《消费者杂志》（*Consumer Reports*）的报道，现在一辆私家车的通常寿命大约是15万英里，也就是说以8年使用期计算，每年平均可以行驶

18,750英里。相比而言，纽约的出租车一年要行驶大约7万公里，那么它们的平均寿命只有3.3年。

如果密歇根大学的研究结果是正确的，而一辆无人驾驶汽车每年要累计行驶20,406英里，那么私家车的使用寿命会更快“用光”，即只用7年的时间就达到了预期的15万英里寿命。将来更为糟糕的情况可能是，废弃的无人驾驶汽车四处散落，垃圾场和家庭后院里堆砌着退役的汽车机身和破旧的引擎。无人驾驶的引入是否会减轻汽车现在施加在我们身上的负面影响，这一点还需要时间来验证。如果无人驾驶汽车反而加大了交通量和汽车行驶里程，那么产生的结果可能是加剧交通堵塞，以及随之而来的空气质量恶化。然而，历史已经告诉我们，新技术绝不会仅仅对从前的状态做片面延伸。我们相信无人驾驶汽车的独特性使其能够避免可能出现的不利局面和对环境的破坏。

如果要求20世纪90年代的因特网立即承载今天的数据通信，那么它一定会卡死在加载过程中。但因特网经历了几年的发展后，很多新增功能使得现在的它可以持续吸收新用户并管理大量增长的数据，这些新增功能包括更高效的压缩技术、更好的光纤电缆，以及更为智能化的路由器。那么，无人驾驶汽车应用可能导致的负面的回弹效应，技术的进步也能适当将其缓解。这种乐观的观点已经得到一些学术研究的支持。

首先，让我们来解决车辆的寿命问题。

根据麦肯锡咨询公司的报告，无人驾驶汽车在刹车和加速方面会更加柔和，这会使汽车的燃油使用量降低15%到20%，同时二氧化碳的排放量每年减少2000万吨到1亿吨。如果麦肯锡的研究是正确的，那么更加平稳的驾

驶行为会使无人驾驶汽车的寿命得以延长。

无人驾驶汽车不仅可以通过上面的方式使用更久，它们还能通过专门的设计来延长寿命。其实，15万英里的使用寿命并不是什么不可攻破的神话。只要有市场需求，汽车公司完全可以设计能轻松行驶几十万英里的无人驾驶汽车。交通部将他们采购的公交车使用寿命预计为至少是12年或25万英里。半挂车则设计成能行驶100万英里，而且它们的引擎实际上可以无间断运转。轨道车辆的寿命甚至可以更长，例如旧金山市的某些海湾快速通行列车（Bay Area Rapid Transit, BART）制造于1968年，时至今日还在运行。

即使无人驾驶汽车的寿命仅与今天的人类驾驶车辆持平，它还是能在现有道路上腾出更多的空间——为了减少风阻，骑自行车的人会彼此靠得很近，这种能量节省策略被称为风拽或者弹弓效应（Drafting）。多亏了传感器、无线通信技术的出现，而且人类也不必再坐在方向盘前，因此无人驾驶汽车在与前后车辆近距离行驶时也能保证足够的安全性。这种车辆结队（Platooning），是指成批的无人驾驶汽车、卡车可以组成紧密的编组队形抵近行驶，进而节省能源。通过减少风阻，更有效地利用道路空间，车辆结队策略能大大节省油耗。

人类驾驶的车辆并不能充分利用路面空间，因为为了安全，我们开车时不得不与前后车辆保持数百英尺的距离，而且在变道时我们也做不到平滑熟练。无人驾驶汽车结队行驶后可以更有效地利用路面空间，让通常会堵塞的地方变得更通畅，这是因为交通拥堵通常发生在公路上的驶入驶出匝道的地方、变换车道的地方，以及交叉路口等地。

得克萨斯州立大学的研究者的一项研究指出，如果美国道路上90%的

车辆改换成自动驾驶的，那么道路通车容量可以翻倍。研究者还预计，紧密的车辆结队行驶能将拥堵导致的时间延误大幅减少，高速公路可以减少60%，郊区道路可以减少15%。由于风阻的影响，卡车尤其易于发生燃油效率低下的问题，如果将自动驾驶的卡车编组，相距不超过3英尺，那么行驶时每辆卡车的油耗将减少15%到20%。

无人驾驶汽车对环境保护的另一个优势，就是重新思考汽车的设计。如果无人驾驶汽车的安全性相比人类驾驶的汽车有了根本性的提高，那么我们就可以对汽车的车体构造做出大幅改进，毕竟它的外形尺寸在一个世纪以来仅做了小幅改进，相比之下，与20世纪初的“原始”汽车相比，现代汽车最大的不同就是增加了更多的“有限的防撞安全装置”。既然无人驾驶汽车的事故率会显著降低，那么新型汽车可以设计得更轻更小，因此也就更加省油。



图1.2 自动出租车应用后，城市重焕活力

来源：Granstudio

取代我们手中工作的交通工具并不只有负责运送人类的无人出租车，包裹和外卖的最后一英里派送也将交由小巧、轻型的自动派送车来实现。多年以来比萨都是美国人的最爱，现如今大学校园里出现了由塑料制造的轮动式自动“比萨机”，这个机器可以利用送餐的10分钟路程刚好完成比萨的烘焙。相比之下，今天大多数地方还在用近一吨重的货车来派送仅一磅重的比萨。而那车一吨的重量绝大部分是为了驾驶员的安全考虑，与比萨无关。

汽车在无人自动化改进过程中一个核心的变化就是它们的驱动能源。无人驾驶汽车更有可能采用电力驱动，但纯电动汽车推广的阻碍之一就是在给汽车电池充电方面缺少能广泛普及的方法措施，不过特斯拉公司正在通过自建充电设施的方式努力克服这一问题。其实，一旦汽车变得智能化能够自动计划路线，并适时地将充电站纳进行程，需要定时充电的引擎也就不会引起太多不确定性。

在无人驾驶汽车给我们个人出行带来更多便利的同时，结队行驶、轻质车身、高效驾驶以及可充电引擎可以将随之而来的某些负面影响降到最低。我们大多数人每天都会做的对环境产生不利影响的另一常见行为就是停车。司机在寻找停车位时常常要苦闷地兜圈子巡视，而无人驾驶汽车可以减去这些麻烦进而提升生活质量，原因在于它们根本无需停车场。

找停车位不再令人头疼

其实，我们很难明确指出一个城市为何令人着迷。

就像是对于艺术由什么构成这个问题的争论一样，一个城市的吸引力来自哪里也不是人们能轻易解释清楚的，但是只要人们身临其境体验过，就能察觉出那个城市的独特韵味。在我们的经验里，那些引得人们去游览、定居并工作的城市多半有着深具活力的城市步行文化。街道上的行人越多，这个城市的热闹氛围也就越有趣，街道上餐厅和商场里的消费交易也就越频繁。

一个城市停车场的设计规划对于城市的个性韵味有着惊人的显著影响。美国东海岸的老一代城市与欧洲的很多城市类似，都是在汽车广泛推广前就规划并开发出来的。这些老城市与汽车广泛普及后才发展起来的新城市有着截然不同的城市氛围。老城市给人的魅力通常被称为步行享受。停车场并不能提升一个城市的魅力，这一点毋庸置疑。

停车经济学家唐纳德·肖普（Donald Shoup）对寻找路边停车位所付出的隐藏成本作出了这样的描述：“堵塞交通、引发事故、浪费燃油、污染空气，还破坏了步行环境。”虽然估计值存在差异，但是美国人为了寻找免费的路边停车位基本上都要花个3.5分钟到14分钟，这对城区的交通造成了极大的拥堵。对于如何改善因寻找车位导致的交通不畅，肖普给出的建议是提高路边停车位的价格。更为有效的改善措施就是干脆取消市中心的路边停车位。

就像血小板过多会阻塞动脉一样，停在路边的车也会阻碍我们的街道交通。平均而言，一辆车的停泊要占用14平方米的人行道。如果你将通向停车位的通道算上，那就会占用更多空间，确切地说单单一个停车位就要占用100平方米大的空间。一般而言，汽车有95%的时间里是停在某地的。这对空间更是极大的浪费。

汽车十分贪婪，大多数汽车都需要不止一个停车位：一个家里的，一个公司的，当汽车的主人下班后去超市或健身房时则还需要一个停车位。这些停车位很少会同时启用，当汽车停在公司里时，家中的停车位就一直空着。

当把时间因素纳入模型时我们发现，为了停车，我们损失了更多的空间。在约翰·迈耶（John Meyer）、约翰·凯因（John Kain）和马丁·沃尔（Martin Wohl）联合完成的经典著作《城市交通问题》（*The Urban Transportation Problem*, 1965）中，他们估算出一辆汽车在寿命期内停泊状态占用的空间是行驶状态下的两倍。起初这个结论看似不合理，但经过一定分析后，它是说得通的。

迈耶、凯因和沃尔的计算核心是“面积小时（Area Hours）”概念，即汽车的所需用地不仅仅是空间的占用，而是空间和时间的结合。平均来看，一辆汽车一天之中有23个小时都处于停泊状态。因此，停泊状态下的汽车所消耗的面积小时，远超使用中的汽车。

那么多宝贵的市中心房产资源怎么能用于存放汽车这种“可耻”行径？在规划城市时，停车场作为不可避免的“灾祸”却被广泛接纳了。尽管我们在建造每一条新公路或新的公共建筑时会分析研讨，但很少有人质

疑停车场存在的必要性。

停车场广泛建设的另一原因是很多城市要求“必须建设”。大多数城市实行严格的分区法规，要求每个区的停车位数量必须达到某一数值。当某些人想要开新餐馆或建设新公寓大楼并向市政府提出申请时，分区法规就会冒出来提要求——想要获得市政府的批准实施，大多数市政法规都会要求建设一定数量的新车位以匹配使用或居住在这个新建筑里的人员。

分区法规要求停车位必须提前做好规划。毕竟，没有人愿意兜着圈子找停车位。然而，几十年来强制实行的“最少停车位建设政策”，导致了未曾预料的后果，那就是现代化的新兴城市被用于存放车辆的无用空间搞得千疮百孔。衡量停车场对于现代化城市观感造成巨大影响的另一种方式，就是计算一座城市的停车场总覆盖率，或者说城市中用于停车场建设的空间总和。城市里，很多停车场是一层叠一层的。设想一下，如果城市中停车场的每一层都展开铺平了，那总共会占用多大面积？

越是新兴的城市，对于停车场的屈服就越严重。洛杉矶市政府就宣称市中心城区有107,441个停车位。如果这些停车场摊在一个共同的二维平面上，它们的表面积累加起来能达到331公顷，这相当于市中心整个城区408公顷占地的81%。花大精力投入停车场建设的另一个代表性城市是澳大利亚的墨尔本，在那里，停车场面积达到了市区总面积的76%。得克萨斯州的休斯敦市，停车场面积则占到了市区总面积的57%。在伦敦和纽约这样的老城市中，停车场占用的市区空间要少些，大约有18%。在停车场建设方面投入空间最少的城市有曼谷（8%）、东京（7%）和马尼拉（2%）。

如果汽车不再需要在闹市区停靠，那么城市的规划者就会发现有一块很

大的未开发空间可以重新计划。这样一来，界定不同种类建筑该在何处建造的城邦法典中，就可以取消对每个新建商圈或住宅楼必须配备一定数量新车位的要求。城市规划者就要忙于一件更为激动人心的事情——将停车场改造成更宜人的空间。不管怎样，结局都会出现一个新的城市乌托邦。

乍看之下，移除市中心的停车场似乎成了一剂特效药，可以一扫很多城区晚间的低迷。将停车场改造成公园、运动场或路边咖啡馆会为弥漫低沉之气的市中心注入活力，还可能带来很多新的工作需求。而且随着人们无须再在城市中找停车位，交通拥堵也会变得更少。

设想下这个城市可以变得多干净、多美丽。没有了停车道，街道立马变成了宽阔的林荫大道。兜来兜去找停车场的车辆变少了，空气也就变得更洁净清新了。乐观派认为无人驾驶汽车将使得所有的城市具备某种活力，即使是洛杉矶也可以。

然而，一个更为理性的观察者会指出，从停车场手中收回大量用地的过程中可能存在其他风险。城市居民和规划者必须共同出谋划策，研究如何将先前的市区停车场转变成同样具有价值的黄金地段。从“停车场变革运动”中受益最大的城市将是那些方案考虑周全的城市，它们的方案会从更宏观、更具发展性的视角考虑无人驾驶汽车对城市景观的影响。

不过，一个城市市区的健康发展依赖很多因素。其中一个核心的问题就是，城市预算经费中有多少是源于违规停车的罚单和停车收费。如果这笔钱停止征收，市政预算会面临怎样的处境？

需要探究的其他因素还包括当地人口的密度和组成。如果市区停车场

被改造成富有吸引力的新型住宅楼和商务区，是否有足够的人愿意到这里安家？这些居民的消费支出是否足够维持新型商业区的流转？最后，如何避免城市郊外因人员流出变成“鬼城”？

将停车场变成有用的居民区和商圈将是一个城市必须解决的难题。便利的个人出行还会带来另一个意料外的结果，即充满活力的市中心可能会逐渐归于沉寂。唐纳德·肖普就有关停车方面的知识出版了一部鸿篇巨著《免费停车的高成本》（*The High Cost of Free Parking*），书中讲了一个警世故事——洛杉矶市政府在市中心一座音乐厅下面直接建造了一个崭新的六层停车场，却遇到了令人意想不到的结果：

1996年时，洛杉矶迪士尼音乐厅刚建好，市政府还面临财务困境。由于坚信便利的停车环境对音乐厅的成功很重要，市政府贷款建造了庞大的地下停车库。这个全新的停车场包含2188个停车位，修建过程花费了市政府1.1亿美元，平均每个停车位花费大约5万美元。洛杉矶的城市规划者声称，一旦停车场完工并投入使用，迪士尼音乐厅地下停车场的高额投入很快就能通过停车收费偿还。这个假设的基础是为了新的地下停车场能大量使用以收回成本，这个新的音乐厅每年举办的演出至少要有128场是满座的。

音乐厅要为停车场服务，而不是反过来，这无疑是个讽刺的现象。而且，这个项目给我们的警示还不只于此。针对直接在市中心黄金地段的地下，花费大量金钱建造超级便利的停车场，洛杉矶市政府吸取的另一个教训是过度高效的停车设施同极为低效的停车设施一样，对于城市的健康发

展都有着致命的影响。

既然停车场就位于音乐厅的正下方，听众从地下停车场可以直接乘坐电梯，抵达靠近座椅的内部通道入口，这待遇很尊贵，有些许皇室的感觉。但是，去音乐厅的人走出汽车后径直进入音乐厅，他们就不再需要在通往音乐厅的街道上漫步。新的音乐厅的确会把人们吸引到洛杉矶市中心，但这些人并不会在附近的餐馆就餐或光顾当地的商场。

对于过度便利的地下停车场所产生的意外结果，肖普进行了具体的描述：“相比洛杉矶市区而言，几乎每个人都偏爱旧金山市区。在洛杉矶，人行道上很少有人，夜晚时甚至可能让人有不安全感。如果每个去音乐厅的人直接开车到那里的地下车库，那么建造这么一个宏大的新音乐厅对于营造城市活力起不到丝毫帮助。”

肖普的观点是对的，旧金山的交响乐厅附近的街道有着更具生机活力的步行者，因为那里的音乐厅停车场较小（只有618个停车位），而且跟音乐厅之间还隔着一小段距离。正是这略微不便的交通状况，迫使去旧金山音乐厅看演出的人必须在市区街道上多花点时间。去音乐厅的人无论是选择公共交通还是停车后走一段距离，他们的出现都会为音乐厅外的街道增添活力和人气。

对于无人驾驶技术而言，洛杉矶交响乐厅给我们的警示是每个便利的新技术都会有隐性成本。无人驾驶汽车的确会减少城市停车场的需求（这是一件好事），但它们的便利性也会降低那些因停车而产生的收益——市区内商店餐馆或商业步行街因行人经过带来的生意和利益（这是一件坏事）。如果无人驾驶汽车能像实施军事打击一样准时定点地将乘客投放到

他们的目的地，那么这种诱人的高效出行就会产生隐性成本——失去人们因徒步出行而给城市街道带来的活力和收益。

现代的城市街道在规划设计时就与停车位紧密不分。而在未来的几十年里，无人驾驶汽车将逐步淡化停车位的作用，重新塑造现代型城市。便利的个人出行会产生另一个附带影响，即人们会寻找新的居住地。

更短的通勤时间和更少的花费

类似孟买、墨西哥城以及上海这样的城市，居住人口超过2000万人。到2050年，预计全球城市人口数量增长将会接近1倍，从目前的33亿人增长到64亿人。这些特大城市的纵深将会横跨几英里，街道上也会塞满各种车辆，因为居民每天不得不外出上班，并且还拼了命地想避免拥堵，最后反而更加拥堵。

随着城市人口的扩张，市政府需要对空间利用做出更明智的规划。在无人驾驶汽车逐渐变得经济实用的大形势下，这种技术对生活改善的一种方式就是将曾经的停车场改造成公园和住房；另一种提升城市生活质量的方式就是提高通勤的效率，降低消耗的时间。

每天，人们都要花费大量的时间用于开车去上班。在美国，单程通勤的平均时间大约是30分钟，一天就要占用超过1小时。通常而言，这一小时多的通勤时间还大多是一个人孤单度过的。

新形式的交通工具会改变人们对距离的感知方式。在20世纪50年代，

汽车的兴起让人们能更加便捷地去城里工作。同纽约城中数万的新潮富人一样，我的祖父母满心欢喜地从曼哈顿市区搬迁到了皇后区的郊外。为了匹配崭新的家，他们买了一辆奥尔兹牌轿车，那辆车的体型堪比坦克，只能勉强塞进他们家的车库。

我的祖父母离开城市有着多方面的原因。虽然房屋中介会告诫买卖房子的关键就是看“地段”，但是当人们在决定到何处定居的时候还是会考虑其他因素。大多数人在选择房子时会考虑房价、房子与工作地点的距离（如上班时间）以及当地学校的教学质量。虽然人们想尽力在工作地点附近定居，但是为了能找到一个他们能负担得起并同时满足其他标准的房子，人们也常常只能住得远些，并忍受那漫长而紧绷的通勤过程。

表1.1 不同大小城市的平均通勤时间大约是半小时

城市片区	人口（人数）	通勤时间（分钟）
纽约-新泽西北区-长岛	18,919,649	34.6
洛杉矶-长滩-圣安娜	12,844,371	28.1
芝加哥-乔利埃特-内伯维尔	9,472,584	30.7
达拉斯-沃斯堡-阿灵顿	6,400,511	26.5
休斯敦-舒格兰-贝城	5,976,470	27.7
费城-卡姆登-威灵顿	5,971,589	28.6
华盛顿-阿灵顿-亚历山大港	5,609,150	33.9
迈阿密-劳德代尔堡-波姆庞帕诺滩	5,578,080	27
亚特兰大-桑迪斯普林斯-玛丽埃塔	5,286,296	30.3
波士顿-剑桥-昆西	4,559,372	28.8

(续表)

城市片区	人口(人数)	通勤时间(分钟)
旧金山-奥克兰-菲力蒙	4,343,381	28.7
底特律-沃伦-利沃尼亚	4,290,722	26.1
里弗赛德-圣贝迪诺-安大略	4,245,005	30.6
凤凰城-梅萨-格兰戴尔	4,209,070	25.8
西雅图-塔科马港-贝尔维尤	3,447,886	26.9
明尼阿波利斯-圣保罗-布鲁明顿	3,285,913	24.8
圣迭戈-卡尔斯巴德-圣马克斯	3,105,115	24.1
圣路易斯	2,814,722	24.8
坦帕-圣彼得堡-克里尔沃特	2,788,151	24.1
巴尔的摩-汤森	2,714,546	30
丹佛-奥罗拉-布隆菲	2,554,569	26.5
匹兹堡	2,357,951	25.9
波特兰-温哥华-希尔斯伯勒	2,232,896	24.9
萨克拉门托-雅顿阿尔卡德-罗斯威尔	2,154,583	26.2
圣安东尼奥-新弗洛尔	2,153,891	24.6
奥兰多-基西米-桑德福	2,139,615	26.3
辛辛那提-米德尔顿	2,132,415	24.2
克利夫兰-伊利里亚-曼托	2,075,540	24.5
堪萨斯城	2,039,766	22.5

来源: <http://oldurbanist.blogspot.com/2012/07/commutes-tradeoffs-and-limits-of-urban.html>

无人驾驶会让上下班路上的过程更便捷，为人们买房提供更多可供选择的地点。随着城市街道不再那么拥堵，居住空间逐渐增加，会有更多的人会搬进新兴起的市区。那些不喜欢城市生活的人也可以选择搬到曾经十分遥远的周边乡村，因为在无人驾驶汽车被大量使用的时候，到那些地方所花费的时间已经大大缩短了。最没吸引力的社区变成了那些“睡房社区”，它们远离市中心，位于市内传统核心功能区的外围。

无人驾驶技术促使人们转变居所的另一种方式，是降低城市居民和小镇居民的平均出行费用。哥伦比亚大学的地球科学研究所开展了一项研究，目的是模拟如果纽约的曼哈顿、密歇根的安娜堡以及佛罗里达州的某个小镇同时推广应用无人驾驶来实现共享出行时，人们的交通出行费用会发生怎样的变化。研究得出的结论是，按需调配的自动驾驶汽车、轻型车体以及因无须购买汽车而省去的大额开销，这些因素汇聚起来的联合效应可以大幅减少个人交通出行费用。研究发现，如果安娜堡的居民不再自购车辆，并改用无人出租车出行，那么他们每英里的出行费用可以减少75%；纽约市民也会从这种搭车中获得很大收益，相比今天人类驾驶的黄色出租车每英里需耗费约4美元，无人出租车每英里只花费50美分；小镇居民出行的费用预计每英里只需46美分。

社交和孤独感

虽然无人驾驶的便利性让人无法抗拒，但有时生活中的低效行为反而会有着积极的影响，例如传统的交通工具能促进社交互动的发生；廉价高

效的个人移动性也可能会产生一些副作用，“孤独”就是其中之一。

根据美国人口统计局的调查，每四个美国人中就有一个独自居住，而且自称“感到孤独”的人群数量每年也在持续增长。为了解释孤独感上升的原因，人们提出了多种可能的理论，有的说是家庭结构变化导致，有的说是工作压力所致，有的说是电视和网络的隔离效应影响，还有的认为社交媒体上蔓延的虚情假意是原因所在。

其实人们感到孤独的一个原因在于他们没有“第三空间”来相处，那样的地方既不能是工作场所也不能是家里，而应该是他们可以闲来无事去转转的地方。过去，教授们可以去教职工俱乐部或茶室来享受这种自在，他们在那里自由交谈，无须遵守什么日程事项，可以闲谈下彼此的研究项目。除了工作以外，人们以前常常会去教堂、保龄球馆或喜欢的酒吧里转转。

现代生活的一个可悲之处就在于它没有为人们提供很多不同类型的第三空间。在北美，很多中产阶级的生活方式会遵循这样一种模式。大学毕业后，不可避免地步入了人生的下一个阶段——成人期。一旦经济条件达到，人们就会购置一个带有私人花园的大房子，而那意味着他们的上下班旅程会漫长而孤单。

交朋友的三个关键因素是：相近的地理位置，反复出现的无目的性社交互动，以及可以让人们放下内心防御的氛围。对于大多数人而言，一天繁忙的工作之后，基本上没剩下多少时间来开展闲暇社交。周末则专门用来打扫房屋花园，或者弥补日常工作日所缺少的“家庭时光”。也许就是因为这繁忙而分离的生活方式，才会使很多中产阶级感到“缺少朋友”。在工作环境中，虽然能够提供邻近的地理位置和反复的社交互动，但那里

并不是让人放下自我防御的好地方。这就是为什么很多人的朋友是在大学时期结交的，而在随后的生活中很少会结交新朋友。大学寝室的同窗之情也是源于生活居所的邻近、反复的接触，以及在那个环境里彼此之间可以放下内心的防御，而展现真实的自我。

同其他技术带来的便利一样，无人驾驶也将提升人们的生活质量——避免人们上下班时所遭遇的烦闷苦恼，并为人们的居所选定提供更多选择。然而，人们为出行便利所付出的代价就是失去传统公共交通中的第三空间。未来，地铁、公交以及火车这些老旧、低效的公共交通将不复存在，而它们施加在乘客身上的人际接触机会也将从新时代居民的手中流失。

不久前，我们在当地报纸上看到了这样一则广告：“去乘公交吧，为了结交朋友。”在推广公共交通方面，这条广告可谓独具慧眼！广告回避了很多人都已知道的事实，那就是公共交通通常没有自驾出行来得迅速或便利。这条广告没有期待人们去相信他永不会认可的东西，而是专注于乘公交车所带来的无可争辩的收益：等公交花费的时间以及因在低效率路线上走走停停所花费的时间，都将有助于你结识新朋友。

事实上，我们自己开车时“浪费”的时间也有着未被发掘的价值。汽车可以作为父母和孩子之间的“移动版第三空间”。近期，一项针对2000名父母开展的调查发现为了接送孩子上学、放学以及参加社交活动，父母们每周平均要投入6小时43分钟。以这个数值计算，父母和孩子每个月“被迫”待在车里共处的时间大约有30小时。

正如每个父母（兼司机）会告诉你的那样，上述驾驶过程中绝大部分时间里是十分单调乏味的。只要无人驾驶汽车变得安全可靠并广泛应用，

很多事务缠身的父母（包括我本人）都愿意将孩子送入无人驾驶汽车，只需挥手送别。然而，如果这种便利性能实现，那么隐藏在它背后的代价就是父母和孩子之间定期交流的宝贵时光会被剥夺。就像智能手机能在餐桌上窃取我们的情感交流时光，无人驾驶汽车会填充上又一个为数不多的技术真空区——卷走我们与孩子交流的空间和时光。

当我回想起曾经在一星期里连续六天早起开车送正处青春期的儿子去参加船员实践活动时，才意识到“汽车时光”的价值。一方面，我讨厌早上5点开始的行程，我的美梦总要被闹钟的钢锤敲击声粗暴打断。如果当时我可以将司机的义务交给一辆无人驾驶汽车，那么我很可能会热切拥抱它的到来。无人驾驶汽车驾驶的安全系数要比我还高，而我还能多享受几小时的宝贵睡眠。但仔细回想会发现，这种便利的解决方案有着一定代价，那就是我和儿子在那段时间通过共处所建立的亲密关系也将随之被剥夺。送儿子去参加活动这件事为我们提供了共处的机会，而且这种情况下只有我们两个人，坐得也近，还没有事务纠缠或其他分心事情影响。在我们开车的路上，儿子会跟我分享他喜欢的音乐。他会谈及自己的观点，而那些观点很多时候是对全球时局的片面分析。无人驾驶汽车的便捷会夺去我们两人的珍贵记忆——当世界还在安静地熟睡时，我们迎着黎明前的刺骨寒风，沿着湖边寻找约定的活动集合地点。

便捷的个人出行可能会让我们变得更加孤独。但事有两面，如果想要外出时只需毫不费力地在你手机上按个按键，那么人们约会、见面的机会就会如同网上交流一样简单。无人驾驶汽车对社交活动的促进方式是它可以变得更具智能化和数据深度化。人们再次乘坐无人移动舱时可以略微关注下“结交朋友”选项，这个选项进行乘客间某些个人特征方面的匹配，

如年龄近似，有着类似的网页浏览习惯，在Facebook上关注了相同的公众人物。

在未来的某天，人们打开浏览器时，也会看到一则关于无人移动座舱的广告：“进来吧，这有个人等待你认识。”

第二章

终极的移动设备

The Ultimate Mobility Device

2014年，谷歌发布新产品的号角声直达“汽车城”底特律。谷歌的最新版无人驾驶汽车原型既没有方向盘也没有刹车。它所传达的信息很明确：未来的汽车就是为全自动化而生的，无须人类驾驶员，或者说不期望由人类来驾驶。

这种设计引发了更大的热议，谷歌并未再按照其前两代无人驾驶汽车的设计思路——改造普锐斯或雷克萨斯等既有品牌车辆——而是对新一代的无人驾驶汽车的车体进行全新的设计，并将各个组件分包给不同的汽车制造供应商来完成。最为难得的是，这个汽车本身就自带一名“专业驾驶员”——其人工智能软件从前两代的处理器中继承了至少130万英里的驾驶经验，相当于已经以每年15,000英里的路程行驶了90年。

几十年来，汽车行业一直在保护性壁垒中运作，借由较高的进入门槛和专属供应商的排他关系，将外界的竞争者挡在门外。随着新一代汽车的吸引力和价值越来越依托软件系统，传统汽车公司将第一次面临汽车行业

外的公司所发起的挑战。谷歌（或者说它的母公司Alphabet）是领跑者，而有消息称苹果公司也正在招募汽车工程师和软件研发人员来建造自己的自动化汽车。近期的一场技术讨论大会上，苹果的副总裁杰夫·威廉姆斯（Jeff Williams）曾将汽车比喻为“终极的移动设备”，更是加深了人们的猜测。

与此同时，许多汽车公司也投入了数十亿美元的资金用于软件研发，并且汽车创新设计的研发中心已经从底特律迁到了硅谷。梅赛德斯-奔驰公司设立在硅谷的分部有将近300名员工，专门负责先进的工程项目设计和用户体验设计。大众公司则投入了工程师、社会学家以及产品设计师等140名研发人员，研究如何将谷歌地图整合到奥迪系列汽车的导航系统中，并持续研发新型的信息娱乐系统。丰田近期也宣布将在未来五年内投入10亿美元用于人工智能研究，并分别在斯坦福大学和麻省理工学院附近，各设立了一个实验室开展研发。

目前，有四种流行趋势，迫使汽车企业不得不重新思考它们的商业模式，这些趋势是电动汽车、无处不在的无线网络、汽车共享以及自动驾驶汽车。随着无人驾驶汽车技术逐渐成熟，这四大趋势将逐渐聚焦、整合为自主化这一个特征。为了生存，汽车企业不得不从自主化运输机器人的角度重新规划他们的产品，而这种定位的调整就需要在员工配置和产品研发过程中做出重大变革。

汽车企业有两个选择：一方面他们可以探索研发自己的软件系统；另一方面他们与提供汽车操作系统的科技公司建立合作伙伴关系，自身只负责制造汽车机体。

汽车和代码

请等一下，现在的汽车不是已经自动化了吗？如今即使一辆廉价的轿车上也会装载复杂的传感设备，而那些设备在十年前只有喷气式战斗机才能安装。今天，一辆普通的汽车用于管理刹车、定速巡航以及动力传输的微处理器也多达上百个。有的软件模块还能在行人突然出现时，或者车辆偏离路线向驾驶员发出报告。实际上，新一代的汽车平均都加载了500万到1000万条代码，这个数字十分惊人。但问题在于代码类型是“错误的”。

汽车软件系统是模块化的，这意味着它们彼此之间大多数时候是独立运行的，也就很少进行信息交换。然而，更为严峻的问题是，今天的汽车缺少适合的人工智能。只有人类驾驶员掌握方向盘时，现代的车载软件才能正常运作，一旦脱离人类的驾驶，这些汽车就几近瘫痪。当今这些依赖人类驾驶的汽车缺少的是一套强大的机器人操作系统，这个系统应该包含必备的人工智能使其能够灵活应对新情况，并能从过往经历中“学习”提升。

想要把汽车驾驶得同人类一样好或者比人类更好，无人驾驶汽车的软件系统必须足够聪明——能够知道自己在哪里，知道周边有什么，能够预计将会发生什么并做出如何应对的计划……但这还不够。除了作出正确的判断——在哪里转向、何时停下、何时刹车以及何时变道等，自动化驾驶汽车的操作系统还必须能监控汽车底层的运动状态，例如告知汽车的人造“肌肉”（制动器）踩刹车，或者对方向进行微调。影响未来汽车行业发展的最大疑问，就在于哪家公司（不管是汽车公司还是软件公司），会首

先研发出一套完善的智能操作系统，而其成功的标志就在于摘下人工智能研究领域的桂冠，拥有准确的人工识别能力。

让汽车公司高管惶恐不安的一个风险因素，就在于未来车辆的“硬件”——汽车的金属框架（底盘）、引擎以及内部装潢等，都必须优先服务于计算机硬件的布置，并且彻底沦为汽车软件的附庸。如果在消费者眼里，汽车的软件系统成为了最具区分意义的挑选特征，汽车公司将失去汽车市场的掌控权。

为了解更多的细节，我们赶到加利福尼亚州的伯林盖姆市（Burlingame）参加了由国际无人车系统协会（Association for Unmanned Vehicle Systems International, AUVSI）组织的“自动车研讨会（Automated Vehicles Symposium）”。汽车行业巨头悉数到场，展示着各自公司最新、最强大的自动化驾驶技术。会议的氛围让人不禁回忆起俱乐部里男人们的扑克牌游戏——玩家之间看似欢声笑语、一片祥和，但表象之下的事实是大家互相之间都不清楚对方的底牌。

无人驾驶汽车领域的玩家都把自己的底牌藏得很严实，这是我们的第一个发现。我们通过邮件联系了五六家汽车公司的相关人员，结果没有一家回应我们的采访请求。当我们抵达谷歌公司专门负责研发无人驾驶汽车的Google X实验室分部，几番提出采访请求无果后，一名行政助理委婉地建议我们可以到邻近的帕洛阿托（Palo Alto）的计算机历史博物馆里看一下无人驾驶技术的发展历史。

我们认识到的第二点是，汽车无愧为陆地霸主。三天里，我们匆忙地往返于会议酒店和住所两地，那是一段惨痛的步行经历——顶着硅谷炎热

的气候在几条十车道的公路上来回穿行。每天早上抵达会议厅时，大家都已是汗流浹背、头晕目眩，但还是庆幸自己已经从烦躁不安的驾驶员的魔掌中逃出，而他们还在那里催动着各自重达两吨的钢铁猛兽，随时可能夺去行人的性命。

第三，我们发现汽车公司和汽车零部件供应商才刚刚涉足机器人操作系统的创建。演讲者轮番登场，用漂亮的设计和精湛的口才展示着各自公司的研究成果，如自动泊车软件能够识别路标的机器视觉技术。可是在对无人驾驶汽车的大肆宣传之下，汽车巨头呈现的所谓的“自动化汽车”项目实质上只是驾驶员辅助系统。

而讽刺的是，汽车公司可是机器人应用领域的专家，只不过是另一种机器人。机器人工人主要应用于汽车行业，它们的手臂无休止地在产线上组装汽车，给汽车喷漆，并制造汽车。衡量一个产业或一个国家对机器人劳动力依赖程度的一个标准就是机器人用工密度，即机器人员工与人类员工的比例。美国汽车行业的机器人用工密度要比其他行业高出许多，基本可以达到每10,000名人类工人都有1100个机器人共同劳作。随着全球汽车行业的机器人用工密度持续以每年27%的高速率增长，汽车公司每年都会增加使用更多的机器人。

虽然汽车的制造过程中已经大范围推广自动化，但是无人驾驶汽车不会将这些生产线上粗犷的、单线条的机器人视作同类。制造车间、装配线以及仓库里那些辛勤工作的机器人既不能走出来，也不能自主化。它们都被固定在那里，并且由人类技术人员精心编写工作程序，使其只是在高度结构化的环境中执行某项特定的任务。如果它们的环境中突然出现一个

抛过来的球，它们也不会偏离预设的工作任务，也不能从过往经历中学习提升。

大动荡

如果汽车公司有能力制定向无人驾驶汽车发展的过渡进程，它们可能会偏爱一条十分缓慢的推进过程：第一个阶段是持续优化驾驶员辅助技术；第二个阶段是在少数高端车型上安装只在特定情境下使用的、自主能力有限的驾驶模块，这些模块大多只应用于高速公路行驶；第三个阶段是这些自主能力有限的模块向下渗透，应用于便宜的车型。

德勤咨询公司认为这种渐进的过程所带来的增值也是缓慢递进的，“因为汽车制造商会先在高端车型上投入新技术，如防抱死系统（ABS）、车身电子稳定系统（ESP）、倒车摄像头以及车载通信系统，然后随着规模经济发挥作用，再向下覆盖低端市场”。然而，这种谨慎的进程可能并不明智。实际上，逐步增加计算机引导的安全技术来帮助人类驾驶员更好地掌控方向盘、刹车以及加速，以这种方式逐步实现对人类驾驶的辅助的方案存在着安全隐患，而这种隐患需要从长期角度考虑人的生命价值和汽车行业的资金回报才能发现。但汽车公司偏爱这种渐进方案的主要原因，就是该方案可以延长他们对汽车行业的主导权。

无人驾驶汽车需要车载智能操作系统以便感知汽车周围环境，分析接收到的信息并做出恰当的反应。想要开发能实现人工智能（尤其是人工

识别)的软件,企业需要具有技能型人才和一定深度的知识资本储备。而汽车公司已经过于习惯制造复杂的机械系统,它们缺少人才储备和知识储备,也缺少在人工智能研究的丛林中进行有效挖掘的运营管理经验。

无人驾驶汽车给汽车行业带来了未知和不确定性。过去一个世纪以来,将汽车直接卖给消费者是个效益不错的生意。然而,如果无人驾驶汽车使消费者更愿意为每次出行单独付费,而不再购买自己的专属汽车,那么也很难再把这些普通的非智能汽车卖给出租无人驾驶汽车的运输公司。如果某一天汽车公司不得不与软件公司建立合作关系来制造无人驾驶汽车,那么这种结合的结果就是汽车公司只能从最终的销售利益中分一小部分。

在一场通宵牌局里,桌面上的赌注会越来越大,同样,现在无人驾驶汽车领域的市场中也蕴藏了巨大的商机。曾任教于密歇根大学、现通用汽车公司总经理的拉里·伯恩斯(Larry Burns)认为,在人们每年驾车出行的3万亿英里(美国境内)行程中潜藏着一座金矿等待发掘。他说:“如果率先进入无人驾驶汽车领域的公司能够从每年3万亿英里的行程中获取10%的份额,并且每英里收取10美分费用,那么它的年收益就能达到300亿美元,大致相当于苹果公司、埃克森美孚石油公司这些世界最赚钱公司的丰收之年的收益。”

汽车公司对于人类主导驾驶的顽固迷恋,在某种程度上可能是一种战略手段,但这背后还是有很多因素影响的。汽车公司需要对消费者的人身安全承担重大责任。过去几十年来,福特平托车事件(The Ford Pinto Case)以及2014年的通用汽车点火开关故障等工程事故让人印象深刻,迫使汽车公司必须优先考虑产品的安全性。安全性是一项重大的责任,这意味着如

果发生问题，一家汽车公司可能要支出数十亿美元召回问题产品、进行危机公关以及处理集体诉讼。除了安全性以外，汽车公司还承载着全球宏观经济稳定的责任。单是在美国，汽车行业及延展产业价值链——汽车租赁、石油公司、汽车经销商、保险公司、媒体以及医疗——可以累加形成2万亿美元的行业价值，这个数值占到了2014年美国国内生产总值（GDP）的11.5%。大型汽车制造商所犯下的每个重大错误都会耗费整个产业价值链的大量资金，并且有损所有参与公司的名誉。

1979年，当时全球第十大公司克莱斯勒公司的CEO李·艾柯卡（Lee Iacocca）向国会提出15亿美元的借贷申请以挽救公司困境，此骇人一举已经载入史册。当被国会问及，为何他这样一个长期倡导自由市场体系的人会寻求政府援助之时，艾柯卡给出的回答是：“我们公司的14万员工及其家属，4700家经销商及其15万员工，19,000家供应商和其25万员工，都依赖着我们公司的正常运营。简单地说，我今天不是为自己发言，而是代表着数十万依靠克莱斯勒的工作来养家糊口的大众。”

正如克莱斯勒公司牵涉到的经济体量所展现的，大型汽车企业持续几十年来充当着国家经济的支柱性产业。然而，由于汽车行业的封闭发展，制造并销售安全可靠、价格实惠的汽车并不是一件容易的生意。几次访谈之后我们发现，汽车公司的高管普遍表示：“行业以外的人根本不了解现今研发一款新车的复杂性，将新技术嵌入车体架构的困难度，而且人们对监管部门的严苛要求和滞后反应也所知甚少。”

即使只是一辆简单的汽车，制造起来也是十分困难的，它需要先进的物流系统、高端的技术和精准的制造工艺，这是一个世纪以来人们在汽车

制造过程中大量宝贵经验的汇聚。为了采购所需的原材料和零部件，每个汽车公司都与全球数千家供应商有着商业往来。一辆汽车上有着大约3万个零部件，既有车门这种大配件，也有用于车子稳固铆合的螺母这些小零件，而且单是将这些零部件组装成一辆车就需要十七八个小时。

讽刺的是，由于传统汽车行业在制造汽车主要硬件上优势过大，最终它们也许只能依靠此类业务维持生计。所以，对于没有汽车制造优势的纯科技公司而言，也就没有承担公众人身安全这样重大责任事务的经验了。而即使是无人驾驶汽车，也仍需要有一个安全节能的汽车车身才能满足精准化管理的需求。当无人驾驶汽车真正走向大众市场的时候，预计那时也将会出现一批新型汽车公司，它们是由软件公司和汽车企业联合组建而成的，这样双方各展所长以实现互补。就在本书写作之时，一些尝试性的联合已经出现，比如谷歌和福特、沃尔沃和微软、通用和Lyft。

实际上在汽车行业里，不同的汽车公司形成紧密的合作关系已经是经过时间验证的一种商业模式。汽车制造的过程本身就涉及层层重叠的供应商网络，以及不同汽车公司之间错综复杂的贸易往来。一家汽车公司为自家车辆打造的一部分配件随后被打包塞到另一家公司的车里，这种情况是很常见的。前不久我们租了一辆房车去长途旅行，旅程中的长期使用让我们对这车的仪表盘非常熟悉。一星期后，当我们登上机场巴士时，看到了一模一样的仪表盘。后来我们查询了房车和巴士的制造公司，发现它们都用的是福特公司的“E系列”平台——福特公司售卖给产业链下游公司一套标准化机柜和仪表盘，而采购者可以利用该平台自行定制车身外观。

迟早有一天，软件公司和汽车公司之间的合伙关系会成为汽车行业

里的惯例。最有可能出现的情形是，汽车公司将它们的“车身框架”平台出售给下游科技公司，由科技公司将它们升级为自动驾驶汽车。这样一种商业模式最终使得投入的双方实现共赢。剩下的问题就是如何向汽车公司证明这种探索性的伙伴关系会取得成功。这种局面不禁让人想起了曾经发生在个人电脑市场领域的竞赛角逐。20世纪80年代早期我购买自己的第一台电脑时，影响购买决策的最重要因素是电脑的硬件。面对数十种电脑品牌，它们有着各自的操作系统和专属软件，以及互不兼容的外设配件。我的朋友们大多选择购买康懋达公司的Commandore 64电脑，但我选择的是Acorn电脑公司的BBC Micro。随着时间的推移，微软公司对个人电脑的硬件市场进行了大洗牌。微软当时研发的DOS操作系统和后续的Windows系统实现了跨硬件运行，这意味它们可以在任何一台IBM兼容版个人电脑上运行。微软为了推广其操作系统的应用价值，还开放了Windows的应用程序接口（API），并鼓励第三方软件供应商研发为个人台式电脑和服务器平台研发相应的应用程序。

今天，软件成了个人电脑购买时的首选因素，硬件则成了随意组合的日用品。与我年轻时的经历截然相反，现在我选择笔记本时会首先考虑它支持的操作系统，硬件供应商的品牌特性则很少会考虑。“软件优先”模式的特例则是苹果公司。当微软公司将它的操作系统适配到大量原始设备的制造商，即不同品牌厂商的电脑中时，苹果公司则坚持同时掌控自家的硬件平台和操作系统。有些人认为正是因为苹果公司对其自身产品进行了全面掌控，才能实现产品设计的大胆创新，并最终诞生出iPhone、iPod以及iPad这样的突破性设计。

今天，汽车公司和以谷歌为代表的科技公司就如同航线交叉的巨型油

轮，双方都在逐渐驶向共同目的地——从新一代的自动化汽车中攫取最大的利益。汽车公司喜欢渐变的发展方式，即研发驾驶员辅助模块，逐渐延长软件系统对方向盘的掌控时间。相反，谷歌的策略则直接指向完全自动化驾驶。如果谷歌先一步到达，它研发的全自动化汽车将先在一些特定的环境中实现商业化运营。一旦这批早期的无人驾驶汽车被证明安全可靠，随后就会进入日常生活的街道实现常规运营，最终引发变革。如果结果真的是这样，那么，微软的模式将再次胜利：随着现在的青少年逐渐成长为未来消费者，汽车的软件将成为最重要的特征，车身车架只不过是后续插曲。

在微软模式中，汽车公司将负责制造性价比高的车身，而将它们推向大众市场的——谷歌或其他类似的软件公司，会为“裸车”装配一套智能操作系统。软件公司也将随之成为一个产品总控中心，负责安装、测试操作系统，并管理为汽车提供各式各样传感器的设备供应商。汽车公司的地位将被降级至“代工厂商（OEM）”，不仅价值无从体现，而且很容易被取代。如果传统汽车公司的业务转向对公（to B）批量销售，微软模式仍将发挥作用。只要私家车市场按照行业预测开始收缩，那么随着掌管无人出租车运营的出租车公司的加入，只会让汽车制造商的地位进一步降低。如果绝大多数新车是卖给运输公司，而非消费者，那么软件公司的销售机会仍将存在，并且产业中心的地位也不会改变。要知道价格才是核心，汽车的实体车身会被当成普通日用品批量卖给运输公司，而批量销售会导致利润空间更加狭窄。

对汽车公司而言，稍好一点的结局就是采取“苹果模式”，那样汽车公司仍能掌控汽车产品研发和销售的全局。将来的消费者中总会有不满足于大众性交通工具的，有些消费者还是想要购买为特定用途定制设计的自

有高价车型，比如移动办公室，或者充当离开家后的专属迷你居所，另外还要镶上鲜明易识别的个人Logo。想要购买这些昂贵专属车型的消费者往往特征鲜明——一掷千金而面不改色，他们通常会有单独的度假别墅，或者去旅游时宁愿包架飞机也不会去坐普通民航客机。由于销售的焦点是整车的质量，既包括硬件也包括软件，那么就仍需维持与软件公司的合作关系。如果苹果模式行得通，汽车行业里的实际生产者还是能够从未来的汽车产业里获取小部分市场份额。

以人类为主导

在无人驾驶汽车领域，偏爱渐进变革方式的不只有汽车公司。美国运输部和汽车工程师学会（Society of Automotive Engineers, SAE）分别给出了各自对向全自动化驾驶演变的发展路线概要。虽然阶段划分上略微有差异，但是它们有着共同的前提：最好的发展路线要以一系列循序渐进的阶段推进——而不是谷歌式的直接跨越——即汽车搭载的“驾驶员辅助”软件只是临时接管驾驶，一旦出现特殊情况就得立即将驾驶权交回人类驾驶员。

我们并不认为渐进式的演变是无人驾驶汽车最好发展路线。很多原因表明，人类和机器人不应该轮流掌握方向盘。然而，还是有些专家认为最优模式就是人类与软件共享方向盘，并且人类要维持主人的地位，软件只提供仆人般的辅助，比如，麻省理工学院的机器人学家大卫·敏德尔（David Mindell）。敏德尔是一位文笔优秀的作家，同时也是当前机器人领域的一名专家和勤于思考的评论员。他指出：“对无人驾驶汽车而言，最大

的挑战将在于自动操作和驾驶员之间的权利交接。”在人机之间微妙博弈的过程中，敏德尔坚信最好的人工智能软件一定要服从于人类领航员或驾驶员。敏德尔声称，研发能够对环境作出恰当反应的智能软件，能帮助人类与机器的正确结合。在他的构想中，人与软件组合后发挥的效力要优于两者各自单打独斗。

在工程师眼里，这种以人机配合为基础的软件其实就是人类为主导（human in the loop）的软件。在一些情况下，将人与电脑结合确实能产生出乎意料的效果。在手术过程中，熟练的外科医生利用机械臂可以达到超出常人的精准操作。当今，民航客机使用的也是人类主导软件，而且很多商业领域和军用领域也是如此。坚持由人做决策的观点自有其独特的吸引力。将人类的最佳能力与机器的最佳能力精心结合在一起，这实在是一个诱人的思维试验，就如同幻想挑选出各个位置的最佳职业足球队员组成一支足球梦之队。机器准确、不知疲倦、善于分析、擅长侦测规律、执行计算以及精准测量；人类擅长做决策，在看似无关的物体或事件之间做联想，从过往经验中学习提升。

理论上，如果将人与一台智能机器组合，至少应该有一位认真负责、技术娴熟的驾驶员。毕竟，在自动化过程中由人类决策的优势，是可能将人和机器的特长发挥到最好。但实际上，只有当各自（人和软件）的责任界定清晰并保持规范一致的情况下，人类的决策才有可能在无人驾驶汽车上发挥作用。不幸的是，汽车行业人员和联邦运输部人员提出的模型里并没有在人和软件之间的责任上作出标准清晰的设定。而他们的建议方案里，在树立维持人类决策的前提时，给出的责任非但不清晰，而且存在责任转嫁。

渐进式演变策略的核心假设在于，当有突发情况出现时，应该有警示或震动提醒驾驶员需立刻坐回驾驶位以处理情况。这种向全自动化驾驶的渐进式前进路线可能看似合理安全，然而事实上，从部分自动化到全自动化的阶段式演变路线存在风险。机器与人可以在某些情况下有效合作，但驾驶并不属于这种情况。驾驶汽车之所以不适合由人类主导，原因之一在于驾驶是单调乏味的。当一件事变得单调乏味时，人们就十分乐意转交机器执行，所以人们会急切地放弃驾驶责任。

在海军参加军官训练时，我就认识到优秀管理的核心原则之一，就是不要将一个关键性任务分给两个人去做，这是一种被称为“责任分散（Split Responsibility）”的典型管理失误。从根本上讲，责任分散引发的问题，在于负责完成任务的不同成员可能会觉得遗漏一些事项也不要紧，他们会理所当然地认为其他人会弥补上。如果双方都不深究，就会导致任务失败。如果人和机器在驾驶过程中存在责任分散，其结果将是灾难性的。

尽管敏德尔推崇人机合作，但他也曾向我们讲述过一个因人机责任分散导致的惨痛事故。

2009年，法国航空447号班机坠入大西洋，葬送了机上228人的性命。事后对飞机上黑匣子分析后发现，坠机的原因既不是恐怖袭击也不是机器故障，问题出在由自动驾驶模式向机组成员转交控制权的空当。在飞行中，飞机的自动驾驶软件被冰覆盖住意外关机了。机组成员还跟往常一样处于困倦的状态中，突然被召回来控制飞机。面对突如其来的紧急情况，驾驶员接二连三犯下致命错误，最终导致飞机一头扎进海里。

敏德尔将447号班机的事故视为“权力切换失败”的案例。谷歌公司则将人类对自动驾驶软件的过度依赖视为“愚蠢的行为”。谷歌公司2015年10月份的月度报告中，对几年前员工在乘坐早期研发出的无人驾驶汽车出行时的一些行为进行了描述。

2012年秋天，谷歌的几名员工被授权可以在上班路上经过高速公路路段时，启动他们研发用的自动驾驶功能。这个方案的预想是由驾驶员开车上高速，并线后把车固定到一条车道后，打开自动驾驶功能。公司最初就提醒每位员工这项技术还属于早期研发阶段，他们需要全程保持对车子和路况的关注。车上还搭载了摄像机以便全程记录乘客和汽车的状态。一次实验过后，员工们对自动驾驶汽车给出了全面的肯定。所有人都在描绘享受到的便利，不用在上下班时间烦闷地困在高速公路上，到家后能精神焕发地与家人共度快乐时光。然而，当工程师团队回看这些车载录像时，问题出现了。大部分人的注意力都从方向盘上移开了，他们做着简单放松活动，享受这片刻的自由。有的人甚至完全离开了驾驶员座椅，爬去后排座椅上找手机充电器。

谷歌公司报告中，我们已经看到了人类的本性，这也是典型的责任分散，或者工程师将其称为“自动化偏见（Automation Bias）”：一旦看到技术有效，人们就会很快地信任技术。结果就是，当人们被鼓励交出驾驶权去享受放松时，他们很难再投入到驾驶任务中了。

因此，谷歌认为无人驾驶汽车的发展过程中是没有中间地带的——

人和机器不应该共同掌控方向盘，这种观点看似有风险，但考虑到消费者的人身安全，这又确实是最谨慎明智的发展路线。自动化会在两方面削弱驾驶员的能力：首先，诱导驾驶员投入另一件事，比如看书或看视频，而这会直接将驾驶员的注意力从道路上引开；其次，损害了驾驶员的情境意识，或者说妨碍驾驶员对驾驶环境中关键因素进行识别，并做出迅速恰当的反应。两个因素结合起来，会导致分心的驾驶员完全不知道车外面发生了什么，这足以清晰地说明在驾驶中责任分散是十分危险的行为。

先进技术能够将人们从乏味任务中释放是一个很大的诱惑，而人们需要一些数据实例来接受这种诱惑。通用汽车公司和美国联邦公路管理局（FHA）资助弗吉尼亚理工大学开展了一项相关研究。研究人员对驾车行驶在测试道路上的12名驾驶员进行了评估。每辆测试汽车都装有两类驾驶员辅助软件：一个用于车道保持辅助系统；另一个是管理加速、减速的自适应巡航控制。研究旨在测试，当具备了能够维持车道、控制车速以及掌控刹车的驾驶员辅助技术后，驾驶员的反应速度是怎样的。为了观察评估驾驶员在研究中的活动情况，每辆车都装有数据收集和记录设备。

12名参加者的年龄在25~34岁之间，均是从密歇根州底特律市的普通市民中招募的，并提供80美元的参加补贴。研究者要求驾驶员假装要做长途旅行，并且对他们带手机进行测试驾驶也不加阻拦，还为其提供触手可及的读物、食物饮料以及娱乐媒体。当参加者进行测试时，研究者会向驾驶员说明还有一名试验人员来做乘客搭乘测试车。另外，还会告知驾驶员试验人员有事情要做，需要在行程中就完成，所以试验人员在行程中大部分的时间里要盯着笔记本电脑里的DVD录像。

给12名参加者设定的场景是在测试跑道上模拟日常的高速公路驾驶，研究人员用设备对他们的反应和行为进行评估记录。研究者的目的有两个：其一，衡量其他事情的诱惑性，比如吃东西、看书或看视频；其二，评估在软件能负责大部分驾驶行为时，测量驾驶员的注意力转移程度。也就是说，研究者想要测试自动化驾驶技术是否会诱发驾驶员作出不安全的驾驶行为，例如精神不集中、握着方向盘时做其他事情，又或者丧失情境意识、失去对驾驶环境中关键因素的识别能力。

同车的试验人员“假装工作”的状态，结合自适应巡航控制软件和车道保持软件的出色发挥，诱导参加测试的人员误以为很安全，他们无须再花精力去专心驾驶。在将近三个小时的测试驾驶行程中，启用了几种不同的自动驾驶辅助技术，结果证明，大多数驾驶员在有技术辅助他们驾驶的时候，都不自觉地犯下了三种以上常见的危险驾驶问题：到后座拿东西，打电话，以及收发邮件。车道保持辅助系统尤其容易诱发驾驶员分心。开启车道保持辅助系统时，多达58%的驾驶员会在开车时看DVD。还有25%的驾驶员会利用空闲找些东西阅读，这些行为将车祸的风险提升了3.4倍。同时，驾驶员的视觉注意力也会下降。当车道保持辅助系统接手方向盘时，驾驶员的注意力就会完全游离。

总体来看，在三小时的行程中，驾驶员的注意力有33%的时间没有集中在道路上。更为危险的是，驾驶员经常会做出长达两秒以上的“目光远眺”，这种可能带来致命危险的行为在整个研究过程中发生了3325次。还算庆幸的是这些危险性的目光远眺总共占用的时间只有8%。

12个人确实是一个较小的试验样本，因此还需更多相关研究。研究

中一个比较乐观的发现是，虽然所有驾驶员在驾驶中都想阅读、吃东西、看电影或者发邮件，但还是有些人能抵抗住诱惑。尽管还需要更多研究支持，但该研究还是发现并非所有的驾驶员都会那么快放弃自己的责任和义务。正如研究者总结的：“本研究发现由于个性和兴趣差异，在驾驶员辅助技术的应用上存在很大个体差异，这意味着自动驾驶系统对驾驶员的影响不能一概而论。”

而在这个过程中无疑还存在一个临界点，使得自动驾驶技术并没有给驾驶员带来安全，反而引发更多危险。设想下在弗吉尼亚理工大学的研究中，如果12名参加者是坐在一辆全自动化驾驶的汽车中行进三小时，那么很有可能他们做其他事情的概率和强度会增大，以致里面的人会完全睡过去或者沉浸到发邮件的过程中。那么当突发事件出现，驾驶权需要快速交接时，完全心不在焉或处在沉睡中的驾驶员根本不可能有能力接管方向盘，这就是全自动化驾驶的危害。

宾夕法尼亚大学也开展过一项研究，研究人员与30名青少年进行了一场开放式讨论，论题是“驾驶过程中年轻驾驶员的手机使用问题”。讨论中出现了两个常见观点：

一个是青少年承认他们知道开车时收发消息的危险，但有时还是会做。而即使是最初声称自己开车时不用手机的人，也极不情愿地承认在等红灯时可能还是会发个消息。另一点就是青少年对于哪些行为算是“开车时发消息”，哪些行为不是，有着自己的一套分类界定系统。例如，他们认为开车时刷下推特就不算“开车时发消息”，自拍也不能算。

让驾驶员和软件系统共同掌控方向盘的另一种风险则在于，不经常使

用的技能会退化。就像是447号班机的飞行员，如果汽车驾驶员有机会躲在方向盘后休息，他们一定不会放过这个机会。可是，如果驾驶员有好几个星期不开车，或者几个月甚至几年不开车，那么当在紧急情况下需要他们接手方向盘时，他们不仅不知道车外发生了什么，驾驶技术也可能早就生疏了。

其他事情的诱导以及人机之间责任分散引发的所谓权力切换问题，都是人机互动模式下存在的严重风险，所以谷歌选择跳过这种渐进式自动化的路线。谷歌在其每月活动报告中发布了有关无人驾驶汽车项目的惊人消息：基于前期在部分自动化驾驶上的试验探索，公司的计划以后的战略只集中于全自动化驾驶技术的实现。

最终，我们的试验得出这样的结论，那就是完全舍弃人工的干预，研发能自动完成从A点到B点全部行程的汽车。每个人都认为让汽车自动驾驶是困难的，这确实是事实。但就当前的情况看，我们认为当驾驶员已经厌倦或疲惫，并且软件还在说“放心，我会控制汽车”的时候，还想让人们把注意力继续放在驾驶上同样十分困难。

根据谷歌的月度活动报告，我们在写作本书时，他们的无人驾驶汽车总共只发生了17次轻微交通事故以及1次备受关注的与公交车低速碰撞事故。在17次轻微事故中，主要责任人都不是无人驾驶汽车，而是另一方的人类驾驶员。可是，2016年2月14日，谷歌汽车与城市公交车发生了首次备受关注的“接触”。与先前的17次小碰撞不同，这次事故的责任方是无人

驾驶汽车的软件，它错误地以为只要自己向前开动，公交车会停下避让。

自从与公交车发生碰撞引发争议后，谷歌汽车后来又发生了类似的故事，讽刺的是这些事故都是因为谷歌汽车“开得太好了”。经过缜密编程的自动驾驶汽车会严格遵守交通规则，而含糊的人类驾驶员则不会那么一丝不苟地遵守法规，总会有些变通。“遵纪守法”的谷歌无人驾驶汽车发生事故的典型场景是：在车道并线时或繁忙的十字路口亮起红灯时右转。缺少耐心的人类驾驶员并不知道无人驾驶汽车会严格遵守道路的车速限制和车道保持法规，因此偶尔就会发生相撞事故。

幸运的是，到目前为止，谷歌无人汽车所发生的事故均没有造成任何伤害。在不久的将来，避免碰撞的最好方式就是教导无人驾驶汽车在驾驶时模仿人类的习惯，灵活变通；但是从长远来看，想要解决人类驾驶员的这些问题，最好的方式就是用耐心的软件系统替换他们，这样才能实现驾驶中的全神贯注。

在这场高风险的、全球性汽车行业博弈当中，胜负还难以预料。如果联邦政府通过法令支持“人类决策”的渐进式路线，那么汽车公司就是赢家，它们将继续掌控汽车行业。另一方面，如果最终法律允许，或者从安全角度出发要求无人驾驶汽车必须完全自动化，那么软件公司就会拔得头筹。

作为数字地图和深度学习领域当之无愧的行业领军者，谷歌公司有着显著的优势。从商业策略上看，谷歌在汽车领域的经验不足反而可能正是其核心优势之一。分析师凯文·鲁特（Kevin Root）撰文说：“不同于汽车制造厂商，即使目前渐进式自动化驾驶新功能成为主流，他们（谷歌）也不会因而损失利润，毕竟他们研发的是全自动化驾驶汽车这一终极版的产

品，而且已经取得了明显的领先优势。”

有一点是明确的。无论无人驾驶汽车的发展路线如何开展，传统汽车行业都不得不发展新的核心竞争力。为了继续参与无人驾驶汽车这一全新领域，汽车公司必须掌握研发人工智能软件这一复杂的技术。几十年来，这一技术的复杂性已经令全球所有顶尖的机器人学家费尽心思。

第三章

独立思维

**A Mind of
Its Own**

20世纪90年代时因特网上曾流传过一个老套的笑话。当时还是微软CEO的比尔·盖茨曾吹嘘，如果让微软来给汽车研发操作系统的话，汽车就会进化成高科技产品，而且1加仑油就能跑1000英里。后来这个玩笑得以延伸：通用汽车的CEO气愤不已，列出了详细的多个反驳观点作为回应，规劝微软还是老老实实在地给电脑做系统，不要染指汽车领域。

通用CEO的反驳观点是，如果让微软来给汽车做操作系统，就会出现以下情况：

1. 汽车会频繁且无征兆地发生故障。这种现象太常见，以至于司机只能接受这种情况，重启汽车，然后继续驾驶。
2. 所有车门都会偶尔莫名其妙地锁住。这时司机想要进入车里只能一边拉动车把手、旋转车钥匙，一边抓着无线电天线。
3. 有时汽车还会完全死机，无法重启，司机只能重新安装一

套系统启动引擎。

4. 每当汽车公司发布新车型的时候，买车的人必须重新学驾照，因为所有控制按钮都做了布局调整。

5. 当城市道路重新规划后，司机需再买辆适配改装路面“操作系统”的新汽车。

6. 汽车只能承载一名乘客，除非司机再多花些钱办理多乘客许可证。

7. 油量灯、水温灯以及报警信号灯会被一个通用的“整车故障”信号灯所替代。

8. 安全气囊在使用前还会问：“你确定要打开吗？”

当然，现如今的Windows操作系统要比那个时候的安全好用得多。但我挺喜欢这个笑话的，因为它充分说明为无人驾驶汽车研发一个高级的操作系统，需要耗费多少先进工程技术。软件系统必须完美无缺，能够流畅管理“庞大的金属盒子”将尊贵的人类乘客从A点运到B点，同时还要避免与其他车辆相撞，避让突然出现的行人和自行车。

科技公司在研发智能软件方面经验老到。汽车公司则是在制造、物流以及政治游说方面能力突出。如果不计较双方各自的优势长短，在研发能够引导汽车行进的智能操作系统方面，科技公司和汽车公司都能作出相应的贡献。

无人驾驶汽车的操作系统

《牛津英语字典》里将“操作系统（operating system）”定义为“支持电脑基础功能运作的软件，例如任务安排、执行应用程序以及控制外设设备”。一辆无人驾驶汽车的操作系统也应有类似的职责担当，即支持汽车的基础功能和高级功能，并对接收到的数据实时回馈。但这还不够，无人驾驶汽车的操作系统必须绝对可靠安全，而且它必须内置高级的人工智能。操作系统必须知道现在汽车在哪里，知道周围有什么，能预期接下来会发生什么并计划该如何作出反应。引导无人驾驶汽车的人工智能操作系统，无论从复杂程度还是从监控广度上都应该优于电脑或智能手机上的操作系统。因此，可以说机器人操作系统的研发目前还处于初级阶段。

迄今为止，还没有哪个机器操作系统敢宣称自身已经完全掌握了三大核心能力：实时反馈、99.999%的可靠性，以及超越人类水平的感知能力。工业机器人虽然有着实时反馈的精确性，然而却不具备人类水平的感知能力。引导民航客机的操作系统是可靠稳定的，但它们却又缺少应对新问题的能力。自动驾驶的客机之所以是一种安全的运输工具，那是因为相比繁忙的城市十字路口，天空上很少出现复杂的意外事件。最后，当某些人工智能软件能模仿人类水平的感知时，如Facebook的面部识别软件或苹果应用程序Siri，它们却又无法实时反馈，或者表现得并不十分稳定可靠。

如果无人驾驶汽车的操作系统表现欠佳，那么付出的代价将不只是低下的工作效率，还会关乎于人的生命。实时反应速度不仅要求能快速反

应，还要做到时机恰到好处。管理无人驾驶汽车的操作系统在监控支配汽车的反应时必须精确到微秒级。当台式机的操作系统启动硬盘时，电脑会宕机半秒左右，这虽然让人心烦，但也还能接受。但如果高速行驶的无人驾驶汽车在执行左转弯命令时快或慢了半秒钟，就可能会导致严重的交通事故。

在这个操作系统的可靠性方面，要求也同样严苛。无人驾驶汽车的操作系统不仅需要防范黑客攻击，还必须进行保险性设计，使其可以在出现硬件或软件的问题后迅速重启。最后，正如我们在前面章节提到的，机器人操作系统必须足够聪明，能够感知周围的环境并作出有针对性的应对。

机器人研发的艺术

为了对研发智能化移动机器人的过程有更深入的了解，我们前往机器人研究的圣地——宾夕法尼亚州匹兹堡市的卡内基·梅隆大学（Carnegie Mellon University）。

几十年来，卡内基·梅隆大学一直是机器人研究和自动驾驶汽车研究的最前沿阵地。威廉·“瑞德”·惠特克教授（Prof. William “Red” Whittaker）是机器人研究领域的传奇人物，他带领卡内基·梅隆大学的塔坦参赛队（Tartan）分别于2004年、2005年和2007年获得了政府资助的DARPA（国防部先进研究项目局，美国国防部的下属研究机构）的挑战赛冠军。有些人认为正是这一赛事的举办加速了无人驾驶汽车的发展。2015年2月，共享出行

公司Uber也急切地加入了自主研发自动化驾驶汽车的大军，将研发中心设在匹兹堡，并从卡内基·梅隆大学的机器人学系、计算机系以及国家机器人工程中心（National Robotics Engineering Center, NREC）招募了四十多人的研发团队。

NREC离大学主校区有1英里，享受联邦政府和企业的多项资金补贴，是卡内基·梅隆大学机器人学院对外的一个运营机构。我们前去拜访之时，Uber的触角还未伸到这里。在为写作此书进行调查时，很幸运的是恰巧我先前的一个学生布莱恩·扎雅克（Brian Zajac）就是NREC的硬件研发人员，他“大方”地带领我们参观了一下机构内部。NREC的使命是基于卡内基·梅隆大学计算机科学家和机器人学家的学术研究成果，研发制造可以应用于实际工作的机器人原型机。NREC研发的很多机器人或自动化车辆原本主要应用于军事领域，尤其是用于灾难复原工作，但真正批量化生产的是应用于工业领域的。

那天早上，我们开车去NREC，刚进停车场就闻到了一股工厂里的味道，这无疑在提醒我们已经离开了学术研究的象牙塔。坐落于阿勒格尼河（Allegheny River）岸边的一栋饱受风雨洗礼的红色砖墙建筑，曾经是车辆维修厂房，现如今就是NREC的主楼。尽管时过境迁，但匹兹堡市曾经辉煌的制造业所遗留下的痕迹还是随处可见。穿过停车场时，我们看到成堆废弃的车轮胎旁还停着几辆生锈的废弃牵引车。与布莱恩在大厅寒暄几句后，我们就开始参观这里展出的NREC的早期项目成果。参观者进入大厅后首先看到的一个作品就是“驮马（Workhorse）”，由惠特克及其他卡内基·梅隆大学的研究者于1979年设计制造的一个灾难恢复机器人，主要用于帮助清理三里岛核电站事故后的废墟。虽然它看起来只是一堆架在六轮

小车上上的不锈钢管，但曾经作出贡献的它业已退役，骄傲地挺立在总服务台旁边。

正式参观开始了，布莱恩领着我们穿过了总服务台后边的一扇门，进入NREC内部大片的办公室、实验室聚集区。随着布莱恩登上一架螺旋楼梯后，我们进入了一个光线良好的、有两层楼高的中央大厅，那里是整个建筑的中心。过去，当这里还是火车修理车间的时候，工程师会将有问题的车头拉到这片空地进行修理。现在，这里成了机器人建造的地方。

大厅的天花板上和墙上挂着几台退役的机器人，其中有一个八足机器人，它在鼎盛时期曾大胆地闯入阿拉斯加州的活火山；另一个退役机器人SensaBot曾被用于监控充满危险的油田。现在它们被几个轨道滑索固定着，在那里见证着下一代机器人的发展变迁。在我们逗留期间，可以看到悬在头顶上的SensaBot在墙上偶尔一上一下地滑动，展现着它是如何像运动员般轻松自如地在台阶上跳上跳下的。

在这个机器人大观园里，我们首先驻足观看的是一个名叫CHIMP（CMU Highly Intelligent Mobile Platform，卡内基·梅隆大学高级智能移动平台，缩写后被命名为“黑猩猩”）的明星作品，它重达443磅，有5英尺高，是一个身形魁梧、浑身泛着红色金属光泽的灾难恢复机器人。为了建造CHIMP，25名软硬件工程师共同工作了一年多。我们的向导布莱恩当时就是那个项目团队的首席硬件工程师之一，他那时刚当爸爸，因此他还打趣对我们说CHIMP其实是他的另一个孩子。

我们发现CHIMP只是DARPA资助的一个持续性大项目的一部分，而那个大项目的目标就是研发能够在对人体有危险的或有毒的环境中独立工作

的灾害应对机器人。三年来，DARPA资助了多所大学及研究实验室里的多个类似项目。

为了检验资助项目的实用性，DARPA在2012年至2015年间连续举办了三届名为“DARPA机器人挑战赛（Robotics Challenges）”的竞赛。来自世界各地的参赛队伍都可以在这个竞赛里展现自己机器人的实力。机器人挑战赛会在专门设计的场地里进行，场地会尽可能模仿自然灾害发生后的情景设计——有着散落在地的混凝土块，顶上压垮的钢筋横梁会挡住房门。参赛的机器人必须在没有人类直接现场监管的情况下完成一系列指定的灾难恢复相关任务，比如关掉金属阀门，在墙上挖洞，清理残骸，爬楼梯，或者使用电动工具。普通人自然可以轻松地完成这些任务，但想要编写人工智能软件来引导机器人的身体动作并进行稳定的人工识别则十分困难，因此对于机器人而言，这些任务十分艰难。

在2013年的第一届DARPA灾难应对竞赛中，参赛的机器人（包括CHIMP）中没有一个能完成所有的指定任务。但近年来，硬件传感器和电脑变得越来越快、越来越小，人工智能软件发展飞快，这些技术的进步不仅加速了无人驾驶汽车的成熟，也有助于灾害恢复机器人的发展——在2015年的DARPA竞赛中，CHIMP和另外两个机器人用了不到一小时就完成了所有任务。CHIMP在全球最顶尖的25个机器人中位列第三，这个名次为NREC的研发者赢得了50万美元的现金奖励。

布莱恩对我们说，现在仍能目睹CHIMP的运动展示，但仅限于一些简单动作。为了准备演示，CHIMP团队的研发人员忙乎了起来，开始启动机器人的操作系统。几个NREC工程师聚集在一块显示屏前，共同注视着一个人

十分复杂的操作界面。与此同时，CHIMP还在原地佝偻地趴着，它那淡定的姿态掩盖了背后复杂的过程——尖端的工程学技术要调动并协调它的肢体动作，软件要将它那伟岸的钢铁之躯连贯运作起来。

CHIMP热身完毕后，我们见证了它机械之躯的“华美”动作。它有两个强有力的类似猿猴的胳膊，可以展开变成履带、收回变成躯干的机器腿，以及精巧的可以自由伸缩的钳形手指。它身上的一系列传感器会将环境中采集的数据及时传递，交由操作系统快速处理。

最后，过了十分钟，CHIMP的操作系统终于启动完成了。突然，它开始移动了。就像人类的运动员要热身一样，黑色的机械手臂伸缩了几下，逐渐地，CHIMP复活了。

CHIMP团队的机器人专家密切地监控着它的动作，飞速在键盘上输入各种指令。当CHIMP的传感器对环境检测完后，操作员给了它一个简单的指令：“前进十英尺。”接收指令后，CHIMP依据周围的环境建立了一个3D数字模型，然后开始计划并执行下一步动作。

计算建模完成后，CHIMP粗壮的机械腿流畅地变形为坦克履带，然后缓慢地向前滚动了几英尺。再后来，它举起手臂掉了个头，顺滑而自然地原路返回到开始的地方。回到起始点后，CHIMP的坦克履带又收回成躯干。演示完成后，它就关机了，再一次安静地趴在那里。

运动中的人工智能

CHIMP是个超强的移动机器人，可以达到奥林匹克运动员的水平。但可惜的是，这些先进的灾害应对机器人——即使如CHIMP这样精心打造的，也不经意间透露出人工智能研究过程中的拦路虎之一：把一台电脑的棋艺调教到世界大师水平是容易的，但教会一个机器人从一堆瓦砾中走过却要困难得多。机器人学家汉斯·莫拉维克（Hans Moravec）将这种看似简单任务自动化过程中遇到的困难进行了简洁的归纳，提出了著名的“莫拉维克悖论（Moravec's Paradox）”。他发现“想让电脑在智力测验或国际象棋上达到成人的水平是相对简单的，但如果想让它达到一岁儿童的感知能力和移动水平，则困难许多，甚至无法实现”。

莫拉维克悖论还展现了一个长期存在于人工智能研究者和机器人学家之间的分歧。几十年来，人工智能研究一直寻求模拟人类智能，研发的软件则力图完成人们通常视为高智力表现的任务，如下棋、解谜或数学计算。与此同时，机器人学家追求的是另一个目标，创造能够观察并理解周围环境的人工生命机体，使其技能水平和流畅性至少不输于人类婴儿。

我们人类感觉简单的活动，如“关阀门、捡起工具以及在墙上凿洞”，其实背后是需要超大量计算能力支撑的。就好比从水面上看鸭子游动感觉很轻松，但实际上它的腿在水下猛烈地拍打，为了让机器人的肢体做出自然的动作，它的软件系统要持续进行大量复杂的计算。结果就是类似CHIMP及其灾难恢复机器人伙伴尽管可以在危害人类的物理环境里工作，但它们的反应

时间非常慢，甚至赶不上一个低级生命体的条件反射弧。

观看DARPA竞赛录像，可以看到机器人在完成任务时的速度很慢，而且节奏不自然。它们的动作之所以缓慢，是因为即使仅仅调动一个的简单动作——例如穿越房间并抓起扳手，相应的软件代码也需要整个机器人系统运行一分多钟才能完成。另一个看似简单的活动——环视房间，检测障碍物并计划出穿行路线，同样也会占用大量的运算资源。想要“感知”周围的环境，机器人的视觉系统要扫描环境，然后中央处理器要处理大量信息流以查找参考模式和有意义的信息。当机器人操作系统确认好周遭环境并设定好一个恰当的反应后，身体不同部分的动作调动还需要额外的数据分析，因此也就耗用更多的时间。

巴西科学家苏珊娜·埃尔库拉诺-乌瑟尔博士（Dr. Suzana Herculano-Houzel）针对引导运动、感知和反应的软件系统为什么会如此耗费资源以及难以研发，给出了一些见解。她研究发现，那些长期被我们视为仅能体现智力低级水平的活动，才是生物大脑真正难以掌握的。之所以普遍被误解，是因为经历了数百万年的进化，我们的大脑对于投放大量认知资源解决这类问题的流程已经超级熟悉了。

埃尔库拉诺-乌瑟尔博士设计了一种技术来检测哺乳动物大脑中的神经元数量。经过数年的潜心研究，她发现人类大脑中包含大约860亿个神经元。然而，与很多脑科学家曾认为的观点不符的是，大量的神经元主要用于监控基本的身体机能和条件反射过程，而非高级的认知功能。

人类的小脑位于脑后的部位，平均含有690亿个神经元，主要负责管理身体的基本功能和运动平衡。相反，埃尔库拉诺-乌瑟尔博士的研究发现剩

下的160多亿个神经元位于大脑皮层——承担所谓高级功能的区域，如自我意识、问题解决以及抽象思维。我们应该从人类神经元的分布状态中认识到，机器人的大脑也应该将大量运算能力贡献于知觉和动作的管理。

在访问NREC那天，看完CHIMP的演示后，我们对机器人操作系统有了更深入的认识。考虑到莫拉维克悖论以及将简单肢体动作自动化所需的计算能力，也就不难理解研发一个用于引导汽车的快速响应、安全可靠、聪明智慧的机器人操作系统来的艰难程度了。另外，加上无人驾驶汽车需要更高的安全性，研发这样的机器人操作系统更是难上加难。

然而，我们在NREC参观的最后一个部门，重新燃起了我们的信心。我们眼前一亮，仿佛在灰暗的土地上发现一台亮绿色的迪尔公司（John Deere）的综合拖拉机。布莱恩给我们介绍了另一位NREC的研究员卡尔·威灵顿（Carl Wellington），他专门研究帮助农业公司实现自动化的农业设备。

现代农场里的工作都有着高度复杂的技术，遵循“精细农业”的实施路线。卡尔告诉我们，农民使用半自动化车辆的历史已经有十多年了。在自动化的早期阶段，农民给他们的拖拉机装上高精度的GPS系统，并使用农场管理软件在地图上标出各自的土地区域。这个时代的农业车辆，还是需要驾驶员的。这时的农用车辆只能够沿着直线自动前进，但还需要人类驾驶员在直线的终点控制车辆转向。然而几年后，引导拖拉机自动转向的商业软件就开始在市面上贩售了。

驾驶的控制权：混合人工智能

在工程学难题列表中，编写引导无人驾驶汽车的操作系统难度应该介于CHIMP和自动拖拉机两者之间。无人驾驶汽车的操作系统横跨了两个存在差异的研究领域。其一是控制工程学（Controls Engineering），专门解决机械零部件协调运行的一个工程学分支；其二则是人工智能研究。

控制工程学要处理复杂系统（如机器人的机械系统）的运行问题，这些系统要通过输入和输出信息来与周围环境进行互动。在给复杂的系统输入指令时，机器人学家会区分出底层控制和上层控制。以无人驾驶汽车为例，底层控制管理汽车内部系统的运行方式，如刹车、加速和转向；上层控制则管理汽车的长期策略计划，如导航和路线规划问题。

当控制工程专注于应用软件管理复杂的系统时，人工智能中相关领域的研究者则努力研发具备智能行为的软件，智能行为这一看似宽泛模糊的定义恰恰反映了这一领域的惊人深度和多样性。人工智能研究的理论基础来源于多个不同的领域，从心理学到语言学，再到统计学都有所贡献。虽然研发具备所谓普通智力水平的软件仍然是这个领域的长期目标，但这不妨碍许多人工智能研究者定向研究某一类问题的领域，例如让工业过程更加高效，或者让汽车能够自动驾驶得更安全。

对人工智能技术的深度解析远非本书所能及。为了简化问题，我们将人工智能技术中丰富多元的内容大致分为两派：自上而下的和自下而上

的，或称为符号型人工智能（symbolic AI）和数据驱动型人工智能（data-driven AI），后者就是渐渐被人熟知的“机器学习”。符号型人工智能会将复杂的情境或任务分解成若干标准的指令或规则条目，然后程序员将这些规则写进软件代码，使计算机能够依据推理和查找功能执行这套逻辑规则。相反，数据驱动型人工智能（或者说机器学习）会应用各种算法对大量数据进行处理，然后利用统计学技术分类、排序，最终再解析这些数据。两种人工智能各有优势，但重要的是根据特定的任务应用最适合的人工智能。

所有人工智能程序都在尝试做的一件事就是将复杂未知的“现实世界”转变成有限数量的逻辑“组块”，以便软件进行后续分析。每个组块，或者说独立的情境，可以被称为一种“状态（State）”。一种状态既可以是指棋盘上棋子的特定组合配置，也可以是一瞬间环境中物体固定下来的特定组合配置，所有可能情境的集合被称为“状态空间（State Space）”。

符号型人工智能技术最适合在小型的状态空间里应用，即所有可能的结果可以事先预期到，并且能按照既定的规则解决。例如，相比繁忙的城市街道，工厂装配线的状态空间数量要少得多。因此，引导工厂内机器人执行有限数量的可能性动作和反应的软件，这种基于规则的符号型人工智能是一种有效的技术路线。

几十年来，符号型人工智能长期占据市场的主导地位。直到20世纪末，随着计算力的提升，传感器将可用数据数量从微小提升到海量，机器学习也从冷门学科进入到人们的视野里，赢得了广泛认同。这种技术路线

的巨大优势，就是它不需要由人类程序员预计到一个场景中的所有可能结果，而传统的符号型人工智能技术却离不开这些。当程序员有了强大计算力的设备和大量训练数据，他就可以编写一个机器学习软件，让机器根据手头的素材“学会”自行应对处理，某些情况下，软件还可以学会应对不熟悉的新情境。汽车的操作系统应该囊括不同类型的人工智能软件，以实现各种控制功能的管理。

谷歌公司的“大眼睛的”自动化驾驶的出现与现代智人（Homo sapiens）的诞生具有相同价值，谷歌的无人驾驶汽车并不完全是谷歌天才研发人员在安保严密的实验室里突然搞出来的。实际上，谷歌现在的无人驾驶汽车是近一个世纪以来人工智能和机器人研究的智慧结晶，它所承载的机器人DNA也是从多个早已停止的研究项目中汲取的，有些项目提供了基本概念，有些贡献了突破性技术。

在人们通常的观念里，有一个不全面的观点，即在很久以前，多个人类物种先后灭绝的过程中，现代智人这一支系因其先天优势而得以保存下来。而现代DNA分析证实这种观点是有问题的。人们曾经认为尼安德特人（Neanderthals）已被其优秀的“表亲种族”所灭绝，但事实是他们的基因仍潜伏在我们身上——近期DNA分析表明欧洲人和亚洲人身上都多多少少携带着一些尼安德特人的基因，这说明人类的演化过程并不像我们曾经以为的那样循序渐进、阶段分明。

引导当前自动化驾驶汽车的软件，其发展渊源也同样存在争议，复杂难辨。谷歌汽车底层控制中用于监控刹车、转向和速度控制的软件，延续了1912年设计的军用机器人原型“战争狗（Dog of War）”的技术路线，可

以说是其后辈的后辈；谷歌汽车上层控制中计划最佳路线的软件，使用的搜索算法也有几十年历史了。谷歌汽车通过对比现在驾驶情境与过往经验来实现学习的能力，有一部分源于机器学习技术，而这一技术的最初提出可追溯到20世纪50年代。

底层控制系统：加速、刹车和转向

无人驾驶汽车底层控制的核心工作是将系统稳定在最佳设定值上。反馈控制设备是一种广泛使用却又鲜为人知的判断性装置，通常用于调节引擎的燃油喷射、生产用机械的电压输入，甚至还可以用于调节房间内温度，使其准确地保持在恒温器设定的数值。反馈控制设备应用背后的逻辑支持就在于它的平衡观念：无论是机械设备、电子设备，还是生物体，一个系统需要尽可能维持在稳定均衡状态中。

自从20世纪80年代以来，汽车工程师将底层控制元件应用于防震、刹车和巡航定速控制等功能中。在无人驾驶汽车上，底层控制所涵盖的范围有些许扩大，加入了汽车核心硬件子系统的管理，按照计算好的路径前进的精准调控，刹车和加速过程中的平滑性控制。底层控制的过程是自动化进行的，而且当信号通过控制器局域网（controller area network，CAN）传递到汽车的计算中心时，系统会迅速作出决策判断。如果底层控制运行良好，无人驾驶汽车上的乘客不会觉察到任何运行状态的变化，唯一能注意到的就是这辆车“自我管理”得很流畅。

为了实现底层控制系统的流畅运行，工业工程师进行了一个多世纪的探索尝试。以现在的标准来看，最早期的反馈控制技术可谓粗糙不堪。在18世纪的蒸汽机时代，就出现了一个叫作调节器（Governor）的金属配件，用于调节从锅炉房中导出的压力强度，进而使得无论载荷如何变化，蒸汽机的运转速度都能保持恒定。这种老式的机械式调节器是利用两个轴承钢球组合旋转构成的精巧金属装置。它里面虽然没有什么人工智能，但是安装上它以后，引擎就获得了至关重要的功能：自动调节（Self-regulation）。

世界上第一台自动驾驶车辆大约是在1912年出现的，发明家约翰·哈蒙德（John Hammond）和本杰明·梅森纳（Benjamin Miessner）装配出一台简单的自动引导型小车，所用的工具只是一个电子回路和一对光感性硒光电管。当车上的光感性电管受到光线照射时，一个底层控制系统会拖拽小车的“方向盘”进行转向，使小车朝向光源方向，以实现两侧感光电管的均衡。设计师给这一粗糙的自动驾驶车辆起了个凶悍的名字——“战争狗”，因为这一设备的研发初衷是应用于军事领域，以帮助美国赢得第一次世界大战。战争狗的设计理念简单明了。在靠近敌人防线的地方投放出去，它能自动去执行破坏性任务，无需人类指引。正如梅森纳所描述的：“当敌人的探照灯扫射到它时，它会立即自动向敌人方向前进。”将爆炸物送到毫无戒备的守夜士兵那里。

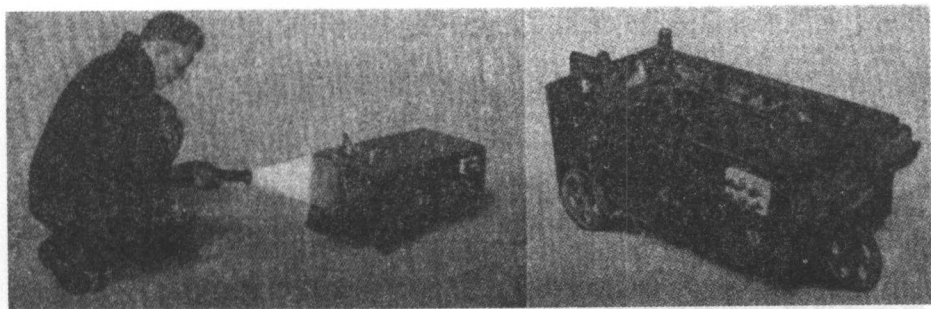


图3.1 战争狗（大约诞生于1912年）

来源：《科学美国人》（*Scientific American*）增刊，June 14, 1919, 376~377页

尽管以今日标准来看，它的设计很是粗糙，但战争狗却是控制工程学的早期应用事例，其背后的科学原理同样适应于无人驾驶汽车。战争狗还是很多现代化杀伤武器的前身，热导飞弹也是遵照同样原理设计的。目前，定速巡航控制系统从本质上讲也是应用了同一套算法逻辑。在车速过低时，自动注入较多汽油来提升车速；如果车速过快，控制器会减少汽油的注入，直到预定速度与实际速度之间的差异降到零。

战争狗身上应用的那些粗糙的底层控制元件已经被淘汰，利用传感器数据调节制动器的电子回路成了主流。现代的底层控制元件都是高等数学的一种拓展应用。底层控制需要使用大量种类各异的算法，以确保某个部件或整个系统的流畅运行。使用较频繁的底层控制算法通常还有其他叫法，有时可以叫控制方案（Control Schemes）或滤波器（Filters），有时被称为反馈控制（Reactive Controls）。随着传感器搜集到的数据越来越多，高级机器学习技术在底层控制技术有着很大的发挥空间。

预测算法通常基于底层控制，以确保机器人行进过程中的流畅性。预测算法可以提升汽车的情境识别能力，方法就是始终留意数字地图的变

化，以便精准计算引擎的燃料注入量，使得汽车可以顺畅、平稳地翻山越岭。有些燃料喷射系统会利用多个不同传感器的数据流，并考虑到多种额外因素，如发动机负荷、空气湿度以及周遭环境的氧气水平，最终计算得出刚好适量的注入燃料以实现发动机的恒速运行。

想要实现汽油发动机的稳定运行，所需应对的重大难题之一就是时间延迟，或者称为“滞后时间（Lag time）”。对于一个发动机或者一个系统而言，时间滞后看似是无关紧要的小问题，但能影响反馈控制设备的运行，即使是最好的设备也不例外。因为给汽油发动机注入燃料的过程，究其核心还是一个机械的、化学的（非电子的）活动，所以因为时间滞后问题产生的难以预料的情况，使无人驾驶汽车的汽油发动机备受关注。这种时间滞后就意味着汽车的启动、加速以及停止都难以实现精准的时间控制。

不过，至少有两种方式可以减少时间滞后问题对自动驾驶汽车的影响：

一种方式是在底层控制中投入更多的计算能力。计算能力强大的电脑可以减弱燃油喷射器所带来的延迟问题，并为汽油发动机的匀速运行和精准计时提供更高的准确度，最终达到较高的平衡状态；解决问题的第二种方式就是更换发动机。

电动式引擎更加容易管理调节，这也是谷歌公司和特斯拉公司不约而同为其原型无人驾驶汽车配置电动引擎的原因之一。只要为电动引擎配置好特定的电压电位，引擎就始终能即时产生对应程度的力矩，推动汽车前进。很遗憾，几十年来大批才智超群的人将他们的智慧用于解决汽油发动机的调节控制问题，因为一旦全电动式引擎应用于无人驾驶汽车，那么那些丰富的智慧成果将变得毫无价值可言。

上层控制系统：路径规划和道路导航

只负责汽车基本运转的底层控制也许可以很快地完成任务，但上层控制系统的工作时间则要长得多，有时可能在整个出行的过程里都需要运行。如果将底层控制系统比喻为“条件反射活动”，那么上层控制系统就类似于“高级心智活动”，即传统上而言的“脑力活动”。上层控制系统的基础是路径规划和导航，而两者都需要通过搜索算法（Search Algorithms）的应用来实现。

搜索算法是基于规则的符号型人工智能的典型应用：如果需要搜索算法迅速解决上层控制中遇到的问题，就必须为这个算法程序提供充足的计算能力。因为大多数问题都有着多个可能解，有多种因素影响最佳路线的制定，所以需要对所有路线进行评估。因此搜索算法会占用大量的系统资源。搜索算法有着广泛的应用领域，现如今它们的覆盖范围可以从国际象棋（检索所有棋子可能性移动后的结果，并进行优劣排列）到导航。为了规划路径，无人驾驶汽车要利用搜索算法罗列出两地之间所有可能的备选路线，并按照优劣等级进行排序。

搜索算法中最为常用的一个是A*算法，是由尼尔斯·尼尔森（Nils Nilsson）及其同事于1968年发明。几乎所有涉及将备选方案排序并筛选最佳答案的问题，都会尝试用A*算法来解决。这个算法为全世界的GPS导航设备提供了智能判断，并且它还应用于当今很多不同领域的软件程序中，如棋牌类对弈和工厂任务调度。

就像阿根廷探戈舞一样，A*算法做的事情也是看似简单实则复杂。在这个算法发明出来之前，早期的人工智能研究者一直困惑于如何让搜索更高效。解决办法就是为这个搜索加入一个巧妙的代价函数（cost-function）。A*算法之所以能提升搜索速度，就在于它使用的代价函数将已经检索路径占用的成本与达到目标位置还需要付出的乐观预计成本相结合，即得出最终需要支付的代价。这个算法的代价函数巧妙地移除了前期搜索过程中的大量重复计算，借助数学方法实现对最短路线的精准搜索。

虽然A*算法是一种通用型搜索算法，但它尤其适用于驾驶活动中的上层控制功能。只要软件研发人员适当调整算法中的代价函数，使其重新运行算法，它就可以用于评估各种各样的驾驶行为需要支付的成本，例如信号灯等待时长、道路拥堵情况、道路维修工程、信号灯数量，甚至需要等待的左转弯的数量都可以计算检索。

尼尔森和他的同事在发明这个算法之初，就开放了其源代码，这一慷慨之举大大加速了数字导航软件的研发进程。自从这个算法诞生以来，它已经成为人工智能研究中最有影响力的算法之一。几年前，我在瑞士的一个小型研讨会上见到了尼尔森，当时大家在庆祝人工智能概念诞生的第50个纪念日。尼尔森目睹了当今人工智能研究领域对每一个细小进展都授予专利的趋势，不无取笑地说：“如果每个用GPS导航的人每次付给我一美分，那么我早就是亿万富翁了。”

今天，上层控制系统和底层控制系统技术都已经比较成熟，并且经过时间的考验后得以广泛应用。然而，它们的成熟却又引发出一个有趣的问题：为什么无人驾驶汽车还没有开始真正在市场中销售？答案就在于莫拉

维克悖论，即看似简单的移动和感知过程，其实十分难以自动化。

正如我们前文提到的，无人驾驶汽车相比其他移动版机器人的一大优势就在于它们是建立在四个轮子上的，无须借助躯干肢体运动。事实上也是如此，由于汽车可以滚动前进，无人驾驶汽车的设计师可以回避类似CHIMP等灾难恢复机器人在自动化过程中所遇到的大量计算难题。但是，莫拉维克悖论的第二个方面，即有关知觉和反应的问题，几十年来仍然难以实现自动化。识别周围环境并做出针对性反应，能够赋予无人驾驶汽车这些能力的软件始终都是机器人操作系统中最关键的部分。

第四章

创建人工识别

Creating Artificial Perception

无论是底层控制系统还是上层控制系统，都无法赋予机器人从视觉上识别周围环境并做出恰当反应的能力。因此，直到最近，人工智能也未能

为汽车提供极为可靠的人工识别。

2015年，黑客乔治·霍兹（George Hotz）发表声明，称只用了一个月的时间，就在自家车库成功改装了一辆无人驾驶汽车。激光雷达、摄像机以及能塞进储物箱内的电脑、网络交换机和GPS传感器，霍兹用这些寻常的设备改装了一辆2016年版的本田讴歌ILX（Acura ILX）。按照霍兹的说法，他的无人驾驶汽车运行状况良好。霍兹还夸口表示，再经过几个月的改进和实践，引导这辆改装版讴歌汽车运行的软件在运作性能上会超越特斯拉S型车（Tesla S）的定速自动行驶模块。

特斯拉的CEO埃隆·马斯克对霍兹的公开挑战并没有置之一笑，而是在特斯拉网站上郑重其事地向霍兹说明了自动驾驶的真正难题：

把机器学习系统做到99%的准确率相对简单，但是在这个基础上再提升0.9999%的准确率却非常难，而这0.9999%才是根本性的需求。去看看每年一度的机器视觉竞赛就能知道其中缘由了，电脑在判断什么东西是狗的问题上准确率可以高达99%以上，但可能偶尔也会把盆栽植物错认为狗。如果在以每小时70英里的速度行驶时犯下这种错误，那么后果将十分严重。

霍兹和马斯克之间的争论，恰恰体现了人工识别软件的价值所在——无人驾驶汽车目前还缺失的核心组件，这还表明了这个组件的发展对于无人驾驶汽车能否真正走向市场有多重要。霍兹成功改装讴歌的事迹，证明了如今一个熟练的开发人员在短时间内便可搭建一辆“不错”的无人驾驶汽车。然而，正如马斯克所说，当需要人们对软件托付生命时，让它的准确率从99%提升到99.9999%才是质的飞跃。

在过去的几年中，移动机器人在定位周边环境方面表现得越来越好。得益于大数据、高分辨率数码相机和更快速的处理器的出现，计算机视觉软件的性能已有了显著改善。另一个催化剂便是机器学习软件的成功应用，解决了机器视觉的一些棘手问题，进而引发了一场人工识别研究的复兴。

无人驾驶汽车技术的最后难关仍然是开发软件来监管汽车的感知和反应。在编写这本书时，我们遇到了一个十分困惑的问题——应该如何称呼这些不能完全归属于底层控制或者上层控制领域的各种新兴软件工具包。思考一番之后，我们索性将这些软件归属到名为“中层控制（Midlevel Controls）”的领域。

在这一章中，对于为汽车提供人工识别和智能反应的各种软件工具，我们用“中层控制软件”这个短语来指代。中层控制软件使得汽车的操作系统能识别传感器数据，感知车辆周围环境的实体布局，并针对周围事物或事件做出最佳反应方案。平均而言，中层控制涉及的事件持续时间——机器人专家所称的“事界（Event Horizon）”——从几秒钟到几分钟不等。相比而言，底层控制运行的事界不到1秒钟，而上层控制运行的事界则需几个小时。

对于人类来说，中层控制活动的一个事例可能相当于从水槽中拿起一个脏的咖啡杯放进洗碗机。诸如CHIMP这样的灾难恢复机器人，中层控制活动的事件可能是视觉传感器发现并确认了地面上存在一个圆形黑色物体。CHIMP的控制软件会将其归类为石头（而不是像影子一样的东西），并且指导机器人的履带安全绕过它。一辆无人驾驶汽车的中层控制活动可能涉及物体形态、状态的识别，即当识别到人行道有个正在骑自行车通过的人时，会让他先行通过，而发现前方道路上突然出现的是个被风吹动的塑料袋时就不会转弯避让。

物体识别的挑战

中层控制软件需要指导无人驾驶汽车，穿行于具有无限可能性的复杂真实情景中。想要体会研发人员所面临的艰巨挑战，只需想象一下编写软件来指导汽车通过一个繁忙的十字路口是何感觉。既然软件应该与一个驾驶员保持相同的标准，那么它就必须遵守车辆管理局制定的一系列标准化交通规则。

减速、准备停车、避让已经在十字路口或刚进入十字路口的机动车辆和行人。而且，还需避让排在前面的机动车辆和自行车，以及在右边道路上同一时间到达十字路口的机动车辆和自行车。

很简单，对吗？但当你将过程步骤分解，问题就变得略微复杂了。第一个任务是编写代码使汽车能够“意识到”即将抵达十字路口，这样它可以“放慢脚步，准备停车”，这一过程可以通过设计GPS定位点和包含十字路口详细信息的数字地图来实现。停车标志或者红灯这样的视觉线索也能起到一定帮助。这一步实际上还没有什么特别难的问题。

现在开始转向下一个任务：避让已经在十字路口或刚进入十字路口的机动车辆和行人。在这一点上，问题开始变得棘手了。首先，什么算是“机动车辆”？软件如何能识别出“行人”这一类群体呢？

为了达到人类驾驶员的驾驶效果，我们假设中层控制软件必须通过某种方式正确地将机动车辆、行人或自行车进行归类。解决这个问题的一种方法是基于规则方法编写一个人工智能程序。为了将无人驾驶汽车可能会遇到的每一种行人、机动车辆和具体交通状况归类，我们必须编写一个尽量详尽的清单，也就是程序员称之为“如果-则编程语句（if-then statements）”。

现在转向下一个问题：你必须通过某种方式描述这些类别，软件才可以识别它们。识别“机动车辆”或“行人”这些分类的一种方式，可能是根据它们看起来更像是什么。例如，机动车辆是“10~20英尺长的矩形物体”，行人是“2英尺宽度以内、2~7英尺高度的两条腿物体”。

对于在过去几十年内一直致力于研究人工智能的研究人员来说，类别

定义方式并非在任何情境下都适用。不可避免地会出现一些特殊情况，让精心构建的最全面的如果-则编程语句失策。例如在十字路口，一个行人可能拄着拐杖，另一个行人可能携带着笨重的包裹，这样他们的总宽度就会增加而超过2英尺；又或者一辆7英尺长的摩托车（因太短而不会被当作汽车）在十字路口呼啸而过，但是软件无法将其识别为机动车辆。

基于规则编写的人工智能软件的核心问题在于，如果没有一个概括性的方法将汽车可能遇到的一切情况进行分类，那么也就无法编写出指导汽车作出反应的全面规则。

由于我们的软件无法每次都准确识别遇到的是汽车还是自行车，所以它也就无法按照道路规则给汽车提供适当反应的建议。这就是研发人员为什么会说这样的话——

如果生活总是条理清晰，那么编写中层控制软件就很简单。

但事实并非如此，在任何一个城市的任何一条繁忙的街道，总是会出现各种新奇的个案超出人们事先的界定。之所以我们总认为感知活动理所当然，再简单不过，是因为人类已经掌握了感知这一复杂的艺术。

Palm公司的创始人杰夫·霍金斯（Jeff Hawkins）在他的著作《智能时代》（*On Intelligence*）里列出了机器感知领域最持久的难题之一：不变性识别（Invariant Representation）。不变性识别是指我们的大脑具备准确辨认同一图像、声音或感觉的能力，即使是在陌生的环境遇到它们或者从一个

新的角度体验它们。不变性识别的一个事例便是我们总是能在一瞬间认出朋友的脸，即使她把金色的头发染成了黑色。

人类的眼睛在向大脑传递数据流时十分流畅自然，以至于我们不需要有意识地拆开视觉场景来理解它。然而电脑通过视觉传感器读取数据流时，则要做大量的加工。视觉数据流在本质上是一个数值数组。机器视觉软件默默地处理这些数值数组，但无法理解数字描述的视觉场景，这是人工智能核心研究的一个难题。

对于人类如何实现几乎完美的场景理解能力这个谜题，几个世纪以来一直困扰着哲学家和科学家。自从第一台计算机被发明出来后，人工智能研究者一直试图用各种技术破解场景理解能力难题，以及与之紧密相关的物体识别问题。有些技术采用逻辑法（如果物体有三个尖角，那么它肯定是……）；有些技术则采用强力破解，即将机器人可能遇到的物体图像尽可能多地存储，并且基于一些比对规则将新遇到的视觉信息与储存的数据库进行比较。这两种方法都会有些作用，但是效率不高，而且无法为机器人提供一项关键能力——在不熟悉的场景中依然准确地识别熟悉场景中见过的目标物。

为了使机器能够自动化地识别物体，开发一款可以从原始数据中提取视觉信息以识别所描述物体的软件是很有必要的。多年来，为了实现这项功能，研究人员已经尝试了各种方法。20世纪60年代，计算机视觉识别软件就开始了其早期开发和研究，其工作原理是从数字图像中提取简单的线条。

这种方法的著名案例是一个名为沙基（Shakey）的机器人，创造者是斯坦福大学研究员查尔斯·罗森（Charles Rosen），他颇为自信地将沙基称为

“第一个电子机器人”。沙基的“身体”由一堆架在小车上的沉重箱子组成，里面装有各种电子设备。箱子的上面是沙基的“脑袋”，一根又高又细的桅杆上架设着相机和电缆。机器人沙基的视觉传感器是一台20世纪70年代的电视摄像机，可以持续采集图像，但生成的图像只有几百像素。

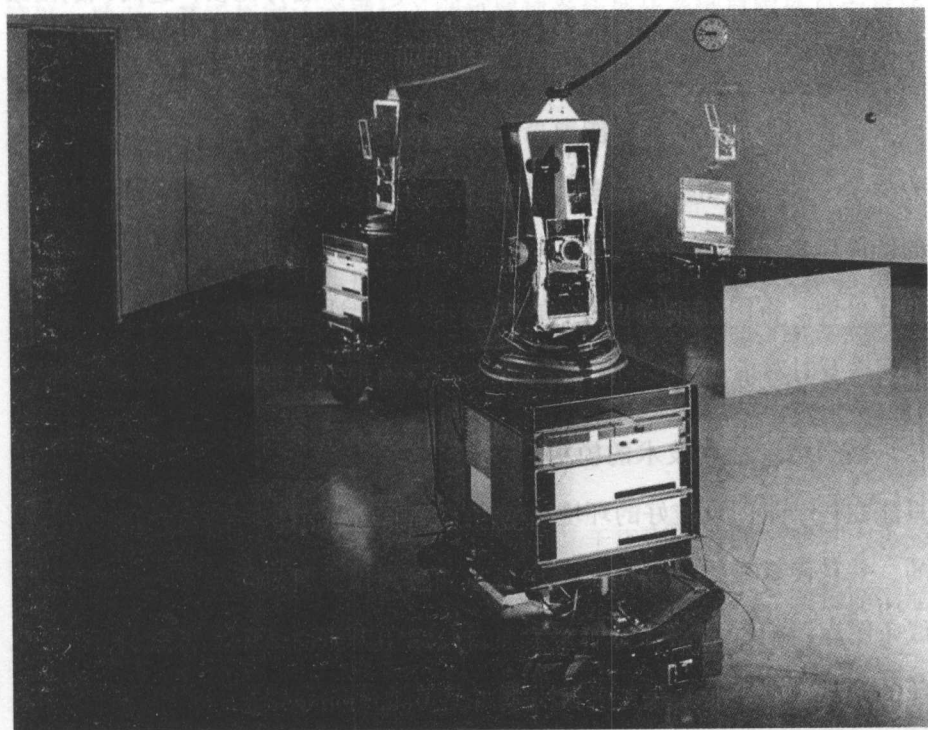


图 4.1 机器人沙基在运动过程中的定格影像（大约摄于1970年）

来源：斯坦福国际研究所

沙基的表现震惊了机器视觉研究界。它对输入其系统内的指令“去找一个红色方块”做出了回应：电视摄像机慢慢转动方向，四处寻找红色方块，当其视觉系统得出结论，确实看到了要找的目标时，沙基将发出任务完成的信号。这在20世纪70年代是一项十分不错的技术。

沙基使用的是一种被称为“边缘检测（Edge Detection）”的人工智能技术，隶属于“模板感知（Template Based Perception）”这一视觉识别技术的大家族。通过分析电视摄像机采集的图像，并将其简化为简单的线条图，沙基可以识别特定形状或颜色的物体。一旦图像被分解，沙基就可以将其与线条图数据库中的锥体、长方体、圆柱体进行比较，进而识别出眼前的这个物体究竟是什么。

边缘检测需要的计算能力、内存和有效的数据存储库相对较小。边缘检测是一种从原始视觉数据中提取有用信息的快速方式，特别是当时沙基的硬件设备还很不成熟，这一特点更为明显。分析一个完整的数字图像要花几分钟。将图像分解成简单的线条画后，让沙基能够在人类还算满意的时间内及时对指令作出反应。

基于边缘检测技术的模板匹配方案，有着显而易见的局限性。沙基只能在一个专门设定好的环境中发挥作用，需要识别的对象只能是它熟悉的，并且没有额外的背景干扰。如果沙基“看到”的对象无法与它储备的线条图库匹配，其软件就会尝试作出最可能的猜测。当在这个环境中投入新的障碍物，比如一个圆柱体、一只猫、飘着的塑料袋，或同事无意中在上午晚些时候漫步走进实验室，沙基的传感器都将捕获这些不熟悉的目标，但中层控制软件却无法识别出到底是什么。

仿照机器人沙基的设计思路，人工智能研究人员研发改善了基于规则的人工智能软件程序，使其可以处理许多不同的数据格式，包括数字图像文件、三维点云（3D point clouds）和视频。有些机器视觉软件可以根据三维深度数据进行深度的测量计算；有些软件专门研究识别数码照片的不同

纹理；还有些程序专门用于识别特定视觉特征的有无与否，并将其与数据库中存取的物体特征进行比对。

许多类似技术得到了很好的应用。事实上，现代工业机器人仍然应用这些技术来执行具体的任务，如检查复杂的电路板或将机器零件分类装进箱里。然而，基于规则（或模板）的机器视觉软件的关键局限性是它最适合结构化的环境，这样机器人的机器视觉只需处理一组内置选定的物体对象。给一个只负责区分螺母和螺栓的工业机器人呈现一只香蕉，它会发现图像库里没有图像能与这个奇怪的黄色物体匹配，就只能干瞪眼了。

当需要为无人驾驶汽车提供情境识别的能力时，基于规则的人工智能研发进展尤其缓慢。在观看2007年DARPA挑战赛时，我目睹了基于规则的人工智能程序在理解交通场景方面的完败（请原谅我使用了这个词语），即因一个失误导致两辆自动驾驶车辆相撞。

早在2007年，计算机视觉软件大多还是基于规则设计的。随着比赛的进行，来自我的母校（康奈尔大学）一辆叫天网（Skynet）的车缓慢谨慎地走在路上，紧随其后的是一辆来自麻省理工学院的汽车塔洛斯（Talos）。突然，康奈尔大学的车无缘无故地停了下来、后退，接着蹒跚前进，然后再次停止，麻省理工学院的汽车试图越过这辆蹒跚前行的车辆。但当后者慢慢要超车时，康奈尔大学的车毫无预兆地又开始移动了，于是塔洛斯撞上了天网的侧面。幸运的是没有人受伤，（后来）两支队伍还都挺享受地成为这次事件的参与者，因为这可能是历史上第一次“无人驾驶汽车对无人驾驶汽车”的事故。对这两辆汽车驾驶日志的分析表明，汽车碰撞是因为它们的机器视觉软件未能出色地完成工作。在事后分析中，康奈尔大学

的团队得出的结论是，汽车突然刹车的原因是“一个测量任务的错误导致天网‘看到’了一个并不存在的障碍物，时不时地堵住了天网的道路”。麻省理工学院团队总结道：“塔洛斯把天网当作了一堆静态物体。”

如果指导麻省理工学院汽车的软件正确地把康奈尔大学汽车识别为移动车辆，事故就不会发生，反之亦然。这个例子充分说明了一件事，即具备识别道路上及其两侧物体的能力对于无人驾驶汽车的中层控制系统是多么的重要。“停着的摩托车”与“行驶中的摩托车”，这两个相似目标之间微小却又关键性的差别或许就会引发人命关天的大事件。

中层控制系统

我们将通过四个组成模块来解释中层控制系统的工作原理：第一个模块是占据栅格的软件工具；第二个模块是一个软件程序，用于识别并标记流入占据栅格的原始数据；第三个模块使用预测性的人工智能软件对障碍物用不确定性锥（Cones of Uncertainty）做替代；最后，第四个模块由一个短期轨迹规划器（Short-term Trajectory Planner）组成，负责引导汽车绕过感知到的障碍，同时保证遵守相应的交通规则。

我们先来谈谈第一个模块，占据栅格。占据栅格是一种软件工具，针对汽车外部环境进行实时、持续更新的三维数字建模。类似于一个包含了数字记录的后端数据库，占据栅格是一个存储了汽车周围实体对象信息的数字存储库。它既可以与中层控制软件的其他模块相配合，也能作为程序

员的视觉参照模型。

流入占据栅格的数据，一些是源于已经存储的高清地图的静态数据，其他数据则来源于汽车视觉传感器的实时导入。随着数据从各种来源流入占据栅格，大多数程序员会使用颜色编码和容易识别的图标，在建立占据栅格模块时代替高频次出现在道路上的物体。

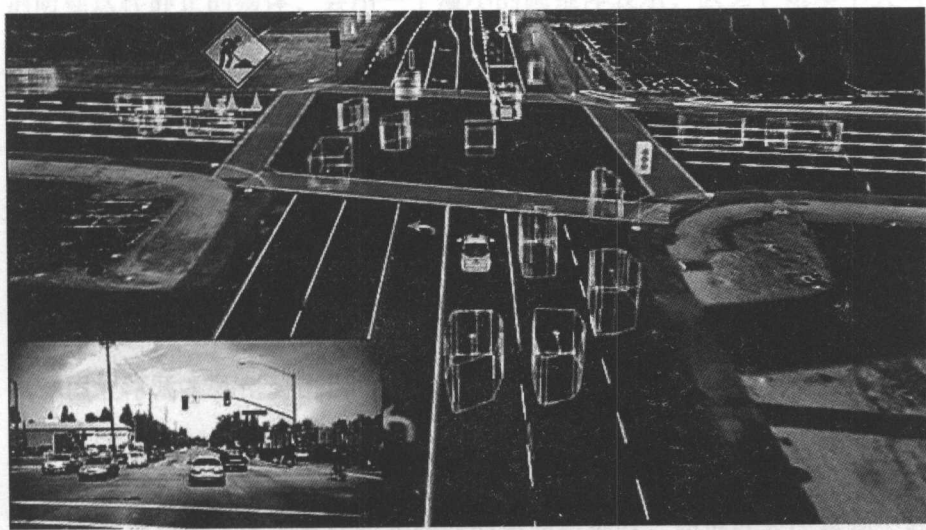


图4.2 以十字路口识别的物体形成的一个占据栅格，同时叠加了高清地图数据和传感器数据

来源：谷歌公司

二十年前，通常的占据栅格是块状显示，而且画面粗糙，视觉感官水平跟老式游戏机中的吃豆人（Pac-Man）一样。在20世纪80年代，那时的无人驾驶汽车配置的是很重的小型计算机和电视摄像机，用于数据建模的占据栅格描绘出的汽车外场景是一帧一帧静态图像断续拼凑而成的。如今的占据栅格能够实时运行，通过列出汽车附近所有物体的滚动更新清单建立一个环境的动态模型。

讲到这里，细心的读者会停下来发现：在占据栅格技术发展史的大部分时间里，软件始终未能胜任正确识别物体的工作。这的确就是事实，但我们需要说明一个新的重要信息：占据栅格只是一个为后续路线规划服务的空间模型。为了解读环境场景，这时需要第二个软件模块来标记汽车传感器导入的原始数据。第二个模块借助深度学习软件对汽车附近的物体进行分类，这样一来占据栅格就可以存储这些信息，以供汽车操作系统的其余模块使用。

通过简化图像中呈现物体的识别过程，深度学习软件终于解开了人工感知的谜题而且促动了无人驾驶汽车的发展。我们将在后文中解释深度学习的内部工作原理，在这里我们将它概括为机器学习软件的一种类型，它使用人工神经网络技术从原始视觉数据流中识别目标。

在过去还无法准确标记数据资料的时代，占据栅格基本派不上用场，因为它只是对周边环境中的大型实物物体的一些粗糙模拟。由于不知道“潜伏”在车外的物体到底是什么，汽车软件系统的其他模块也就无法计算出最佳的应对措施，也无法预测这些“不明物体”下一步要做什么。直到最近，无人驾驶汽车也还只能在几乎没有移动物体的静态环境中稳定运行，例如工厂、矿山、农场，以及沙漠。在这些静态的环境中，无人驾驶汽车的引导软件能够很好地发挥作用，是因为它只需要躲避遇到的每一个障碍，而无须区分它们是什么。

深度学习使另一个技术的应用（不确定性锥）成为可能，该技术也大大提高了无人驾驶汽车软件在动态环境中的性能。这个被称为“不确定性锥”的工具是中层控制软件的第三个模块，用于预测汽车附近物体的位置

和移动速度。一旦深度学习模块标记了一个物体，占据栅格就会显示出它的存在，不确定性锥就会预测物体下一步的运动方向。

不确定性锥为无人驾驶汽车提供了人工智能版的场景理解能力。当人类司机看到行人站得离汽车太近，他就会在脑海中思考要转向避开；在无人驾驶汽车中，利用不确定性锥技术也会进行类似的“脑海思考”。像是消防栓这样静止的物体，会用一个瘦小的圆锥体表示，因为它基本不大可能会移动。相比之下，快速移动的物体会用一个宽大的圆锥体表示，因为它可能运动到的地方比较多，所以它将来的位置是不确定的。

人类驾驶员并不会在脑海中将附近的每一个物体清晰地标记成椭圆锥体。然而，我们潜意识做的处理过程实际上是大致相同的。我们的大脑不断记录更新着周围出现的人和物体，结合以往的经验 and 眼前事物的状态，我们能猜测出这些周边事物的意图并预测出它们下一步会做什么。

机器人专家从天气预报领域借用了圆锥体的概念：如果你曾经看过电视上天气预报中气象学家说明龙卷风的运动轨迹时，你也许会注意到不确定性锥。圆锥的顶端标记的是当前已知的龙卷风位置。锥体尾端宽的部分是龙卷风的可能行进路线，即接下来的几天里它可能会肆虐的地方。锥体的开角越宽，龙卷风最后的终点越不确定。

中层控制软件按照如下方法创建圆锥体：首先，想象下在纸上画出一个物体，在物体周围画一个小圆圈，我们称它为“当前活动圈（current circle）”；然后，再画一个大圆圈，标记出未来十秒钟后物体可能会到达的所有位置，我们叫它“未来活动圈（future circle）”。最后，用两条线把小圆和大圆的边缘连接起来。这就是不确定性锥。

不确定性锥替代了人类驾驶员与行人之间的眼神交流作用。从无人驾驶汽车的视角来看，一个站在路边面向街道的行人会用稍微向前倾斜的锥体表示，表明她随时可能穿过街道。如果她的眼睛不是盯着前方，而是盯着手机，她的锥体图标则是另一种形状，或许更加窄小，因为她并没有准备好继续前进。如果她扫视了一眼无人驾驶汽车，她的锥体图标将进一步缩小，因为汽车的软件会识别到她看见了这辆车，也就不太可能挡在汽车的前进路线上。越不可预知的行人，锥体的形状就越大。摇摆不定的自行车骑行者比静止的行人有更大的不确定性，相应的锥体也就更大。四处乱撞的小狗或追着球跑的孩子，也都会用更大的锥体表示。

有时，即使一个静态的目标也可能会用一个大号锥体表示其不确定性——具有遮蔽性的建筑——虽然它们本身不大可能移动，但是可能会遮蔽了移动的物体。对于死胡同、转弯处，或随时可能会有乘客下车的一辆停在路边敞开车门的汽车，无人驾驶汽车的中层软件系统都会标记一个大号的不确定性锥。静止的校车也可能会产生不确定的大圆锥，虽然校车本身或许不动，但是随时都可能有孩子从车后跑出来。

当前三个模块完成了它们的工作后，第四个模块短期轨迹规划器就可以上场了。当汽车附近的物体都被标记并计算成了大小不一的不确定性锥，无人驾驶汽车的轨迹规划器就能制定出汽车前进的最佳路线。轨迹规划器使用精密完善的算法计算出最有效的前进路线，并保证遵守交通规则，减少行程时间和碰撞风险。

计算机尤其擅长计算这种非线性轨迹预测。在过去的几年中，软件已经被提升到了新的高度——计算机能比人更好地预测出物体的运动轨迹。

特定条件下，可能的结果越多，计算机在计算各种可能性时表现得越好。添加的影响变量越多，如行人可能的运动范围，轨迹规划器的计算结果反而会提升更多。

以目前中层软件的优化升级速度来看，过不了多久，无人驾驶汽车的操作系统将实现驾驶过程三位一体的完美掌握——实时响应，稳定可靠以及智能处理。这是否意味着应该允许无人驾驶汽车上路了？可能吧。但首先，人类乘客还需要无人驾驶汽车拿出更确凿的可靠性证据，以及对机器人司机可靠性的清晰界定。

绝对安全可靠

我们已经讨论过当前的技术瓶颈，操作系统目前仍难以做到完全的可靠性。此外，足够全面的法律标准也需要进一步定义并量化。很多人坚持认为，无人驾驶汽车只有达到100%完美的可靠性才能合法，那意味着没有发生任何碰撞、事故或失误。然而事实是，如果需要完美的可靠性，无人驾驶汽车将永远不会获得合法地位，因为没有任何操作系统能做到始终完美可靠。

操作系统出错的概率很大，而且即使设计最精密的操作系统也会时不时地出问题。在个人电脑上，用户可能会安装一个有漏洞的新应用程序，以至于影响系统的核心程序。系统可能会被恶意的电脑病毒感染；一个新的硬件外设也可能会不经意降低整个系统的性能。

电脑的操作系统有两个特质，使其无法做到绝对可靠和安全：第一，它们包含数百万行代码，数量之大以至于熟练的程序员也很难迅速搜寻到潜在的漏洞；第二，操作系统要承受所谓的故障隔离（Fault Isolation）之苦。类似windows、iOS、Linux这样的操作系统，是以这样一种方式建立的：它们的子部件之间或程序之间几乎没有什么隔离。Windows XP的架构包含大约500万行代码，而这些代码又负责在单片数字“工作区”或内核运行着成千上万个程序。这些成千上万的程序联系还都连接着一个单一的二进制程序。类似于所有登山者都沿着一根绳子前进，如果一个人倒下去，所有人都会被拖倒。

硬件问题经常是计算机操作系统故障的根源之一。机器人操作系统需要管理更庞杂的硬件组件。如果像鼠标、键盘和耳机这些简单的外设都能导致计算机操作系统崩溃，那么汽车“硬件”带来的不可靠性也就可想而知了，比如它的轮胎、刹车、方向盘，以及驱动无人驾驶汽车所需的各种传感器。

每个附加的硬件都需要一个专门的软件驱动程序，使该硬件能与安装在操作系统上的其他部分沟通。所以，硬件的驱动问题是系统故障的另一个更主要的原因。驱动问题可以带来严重危害：操作系统故障中大约有70%是由驱动引发的崩溃，这个错误率是软件代码漏洞引发问题的3~7倍。

操作系统越复杂，越难预测它会出现何种问题。当代码行数达到数千万级时，即使编写无误也可能会崩溃，更何况若有黑客给无人驾驶汽车安装流氓硬件或胡乱修补操作系统，更是会威胁到数据安全。

对于无人驾驶汽车而言，操作系统崩溃引发的风险也更高。当个人

电脑全面崩溃时，最多也就是让人恼火，没有用户会死于“蓝屏”“白屏”“黑屏”。然而，在无人驾驶汽车上，某个硬件的组件停止响应，整个操作系统受困于某种系统性问题并导致运行十分缓慢，这些问题虽然不是灾难性的，但也是十分危险的故障，而系统崩溃几乎就等同于死亡。

现实情况是，无人驾驶汽车肯定会出现软件故障，但还有待明确的问题是，故障率降低到多少是可以接受的。在服务器操作系统建设完善的情况下，系统管理员可以对系统的停机时间和每年因为系统故障导致服务器离线的小时数进行科学计算和记录。停机时间可以是计划内的，如有时需要升级，也可以是计划外的，比如发生灾难性的系统故障。不过值得一提的是，无论是计划内的还是计划外的停机，都会给企业增加成本。

成熟的计算机操作系统得益于系统管理员研发出的提高系统稳定性的一些技术。因此，在过去的几年里，每年服务器停机小时数急剧下降。如今，通常配置齐全的Windows或Linux服务器每年仅会出现几分钟的停机时间。如果对于高度依赖服务器集群的商业世界，一年几分钟的停机时间是可以接受的，那么对于无人驾驶汽车来说这个时长OK吗？

有些人已经厌倦了人类驾驶员由于分心、醉酒驾驶或情绪冲动而造成的大量车祸伤亡，他们可能会倾向于接受可靠性还未达到完美的无人驾驶汽车。只要操作系统可以比人类驾驶得好，这些实用主义者就可能认可无人驾驶车的合法性。理性地说，这样的可靠性标准还算是合情合理的。然而，难点就在于让那些有话语权的反对者给无人驾驶汽车一个机会。

人们对机器设定的性能标准通常要比对完成同种工作的人员要求更高。我们担心如果只是将无人驾驶汽车的可靠性标准设定到人类水准，可

能很快就会出问题。我们对人类驾驶员造成的悲剧已司空见惯。而由机器人司机引起的第一个悲剧事故将会引起公愤，进而让公众在很长一段时间里排斥无人驾驶汽车。让我们提出一个醒目且直接的基线标准：无人驾驶汽车的安全性必须达到人类司机平均水平的两倍。我们可以根据这个基线设定一些数字标准。根据汽车保险行业数据，平均来说，每个美国的驾驶员每17年会有一次事故；从另一种角度看，平均每个驾驶员每开车19万英里就将申请一次事故保险理赔。虽然事故很常见，但引起死亡的事故很少，只占到0.3%。换句话说，人类司机造成的事故种类既有轻微刮擦碰也有死亡碰撞，基本上每驾驶19万英里发生一次。我们就采用这个度量标准，再四舍五入之后——人类：大约行驶每20万英里；事故：1次。

机器人的可靠性可以进行量化，标准就是它在无人工干预的情况下可以单独运行的时间长度。这个度量标准被称为平均故障间隔时间（mean time between failures, MTBF）。像人类一样，所有的机器人有时也会需要帮助。例如，在家里，我们的Roomba（iRobot公司的智能扫地机器人）平均每10个小时会需要被“拯救”一次，通常情况是它把自己堵在椅子腿中间或者把轮子卡在缠绕的电线里。高端工业机器人的平均故障间隔时间会很长，特别是当它们在远程监控环境下操作时，人类的监督人员每隔一个多月才过来检查一次。

如果无人驾驶汽车的安全性需要达到人类驾驶员平均水平的两倍，那么意味着它是人类无事故安全驾驶的平均里程数的两倍，它们发生故障的可能性是每行驶40万英里出现一次意外。由于我们讨论的是里程数而不是驾驶时间与故障的关系，让我们将度量标准称为平均故障间隔距离（mean distance between failures, MDBF）。度量标准会以常见的道路条件、交通条

件和天气条件为基础。让我们来计算一下无人驾驶汽车需要多长时间才能打破40万英里的平均故障间隔距离纪录。如果自动驾驶汽车每天可以行驶1000英里，那么只需400天就可以验证一次平均故障间隔距离是否达到，也就是一年多点。或者说，1000辆无人驾驶汽车可以在24小时内验证一次40万英里的平均故障间隔距离测试。为增加统计显著性，它们可能需要重复测试多次。

教导机器人如何思考的一大优势就在于它们有“蜂巢思维（Hive Mind）”。如果一个机器人学会了某件事，那个软件就可以复制到其他几十个机器人身上，那些机器人就可以用这些知识继续进一步的学习。当有很多不同的机器人系统在一起学习时，它们每个的学习成果可以集中到一个中央知识库，然后共享给每个独立机器人的思维数据库，使得每一个都能更快速地学习。有人称这种集体学习方式为组队学习。

无人驾驶汽车技术的一大优势是，汽车可以以幂次方效率快速学习。无人驾驶汽车车队可以通过吸收彼此过去的行驶数据库进行学习，这正是谷歌、沃尔沃和特斯拉目前使用的一种磨炼其无人驾驶汽车技术的方法。

组队学习需要重复，很多很多的重复。在2014年4月发布的一段视频中，谷歌自动驾驶汽车项目的测试驾驶员普里西拉·诺克斯（Priscilla Knox）说：“我们工作的很大一部分就是走到世界各地，发现汽车可能会遇到的所有各种场景，然后再协助工程师教导无人驾驶汽车如何应对各种情况。”视频文章发布后补充的文字说明是：“由于经历了成千上万种不同情况，所以我们建立了可能性软件模型，按照最可能的情况（汽车在红灯前停住）到最不可能的情况（闯红灯）进行标记。待解决的问题仍有很多，

包括在拿下另一个城镇之前指导汽车可以在山景城（Mountain View）的街道上积累更多驾驶经验。”

比人类驾驶安全两倍

随着存储知识库中驾驶情况的持续成倍增长，无人驾驶汽车将变得越来越能干。然而，为了获得整个社会 and 法律的认可，无人驾驶汽车仍需要一套透明的可靠性标准，一个定义明确的公共标准——规定了安全的和可接受的平均故障间隔距离。理想情况下，无人驾驶汽车的法规需要由制造企业将与政府合作，共同定义出无人驾驶汽车的平均故障间隔距离应该达到多少才是可接受的。

在无人驾驶的新时代，平均故障间隔距离的水平将如同汽车的马力一样，成为一辆新车面向消费者宣传时必不可少的信息。当消费者决定买什么车时，无人驾驶汽车的平均故障间隔距离水平将成为汽车宣传报道的典型特色之一。然而，平均故障间隔距离这个术语有点烦琐，我们建议使用一个简单的术语去描述。

无人驾驶汽车的平均故障间隔距离应该用人类安全水平来测量。毕竟，早期汽车不是用瓦特来表示引擎功率，而是从大家都熟悉的马车能力来衡量——马力。无人驾驶汽车应该用同样的方法来展现它们的驾驶安全性，即我们已经很熟悉的人类在驾驶中的安全系数（humansafe rating）。下面来讲讲如何操作：一辆无人驾驶汽车如果无事故行驶的里程数两倍于人

类的水平，则可以宣称“人类安全系数2.0”。一辆无人驾驶汽车的安全程度如果3.5倍于人类驾驶汽车，则可称为“人类安全系数3.5”；以此类推。无人驾驶汽车的人类安全系数水平，将取决于汽车的软件和计算能力，以及硬件传感器的数量和类型。

从政府层面对人类安全水平建立监管系统，对于无人驾驶汽车行业的每个参与者都是有益的。一辆无人驾驶汽车的人类安全水平将是其市场吸引力的核心部分，汽车公司可以生产专门具有超高人类安全级别的“超级安全”汽车，并从这额外的安全性能提升中收取费用。从立法上讲，也许会要求运送儿童的无人驾驶汽车需要具备10.0的安全系数，但货物车辆具有低点的安全系数也可以上路。对于各种款式的流线型跑车和装有特殊艺术类人工智能及传感器套装的无人驾驶汽车，消费者能津津乐道地向他人谈论其中的安全系数和马力的兼顾的话题。

实际上，一些安全性法规方面的工作已经完成了。有这样一个机器人操作系统已经成功地解决了大部分此类问题，那就是民航飞机上的操作系统。在所有运输系统中，具有最高级别平均故障间隔时间的运输系统之一就是民航飞机。平均而言，每200万次的飞行才会出现一次致命的系统故障（不考虑飞行员失误、恶意破坏、恐怖主义和其他外部因素，事实上80%以上的飞行事故是由那些因素导致的）。

相比过去飞机的安全水平而言，每200万次飞行出现一次故障可以说是极大地提升了安全性，比20世纪50年代的事故率改善了大约100倍。更何况，现在的飞机比那个时候的飞机的飞行时间更长、距离更远、速度更快，这样一个近乎完美的安全纪录更是令人赞叹不已。不管怎样，在过去

50年里，航空电子系统的故障发生次数已经下降了两个量级。

无人驾驶汽车可以从民航飞机学到什么呢？首先是关键的政府监管。现在的飞机要接受更严密的审查，政府对飞机维护和飞行员训练都进行严格的监管。无人驾驶汽车也同样需要政府的监管，但首先应对机器人有充分了解并基于安全数据做出指导。对无人驾驶汽车进行管理时，应有一套合理而透明的机制来权衡并量化机器人驾驶员的能力，这一点是至关重要的。

然而监管只是难题的一部分，更好的操作系统设计同样必不可少。你应该还记得，电脑操作系统的一个致命弱点——目前软件架构的设计思路是所有的系统进程运行在一个单片工作区。将任何东西注入这个软件集群，都会破坏系统的可靠性和安全性，产生漏洞。正如一份评价操作系统安全性的报告所言：“目前的操作系统就像未发明分隔舱之前的船一样，每个漏洞都可能导致沉船。”

无人驾驶汽车需要高度模块化和冗余设计（Redundant）的操作系统。飞机以冗余架构设计而出名，而这种设计的大部分框架都是由美国联邦航空管理局（Federal Aviation Administration）严格管制的。在飞机上，关键的物理机械系统必须有双重或三重冗余设计。例如，民航飞机的燃油管需要双层护套，这样内层的油管发生泄露时，外层的油管就能承接泄漏并侦测汇报。飞机的关键软件子系统也是高度模块化和冗余设计的。航空电子操作系统是由几个相互关联但又各自独立的电子子系统组成。分离是关键，例如，控制引擎的软件与控制起落架的软件是分离的，而且它们都和乘客使用的资讯娱乐系统软件相分离。

故障防护也很重要。每个航空电子系统都是依此设计的，这样系统之

间可以相互检测故障，而且一旦故障发生，也能适当接受。如果多个子系统“产生分歧”，会通过投票的方式来解决僵局。

像飞机一样，无人驾驶汽车应具备一套冗余设计的实时操作系统，这套系统还应内置独立的自我检测系统，而这些都应进行立法规定。例如，一辆无人驾驶汽车应该具备三套独立的视觉感知子系统，每一套系统都有着各自的视觉传感器和独特的技术来分析传感器收集到的数据；汽车的操作系统需要定期校正以确保它运行的是经过最新的数据库训练过的最新软件程序，而且与软件系统匹配的硬件（机械）系统也应该是准确无误的。

同飞机一样，无人驾驶汽车的线缆和车载电脑应该隔离人体，避免乘客无意间地触碰或恶意劫车者有意破坏；汽车的自动驾驶软件应包含强大的监督软件模块，要由该套软件制造商以外的第三方独立机构进行周期性强制检查。如果子系统之间有意见分歧，监管软件要么“调整”系统的性能，要么立即驱动车辆寻求维护。

几十年的人工智能与控制工程研究形成了现代无人驾驶汽车的操作系统和操纵装置。无人驾驶汽车的操作系统需要既可靠又聪明，操作系统必须实时“掌控环境”，并且绝不会发生碰撞。考虑到这些挑战，有些人会疑虑无人驾驶汽车是否真的能被推出。他们会问：“真的会有操作系统能像人一样掌控汽车吗？”更有甚者提出：“为什么直到现在，才开始发明出无人驾驶汽车？”

第五章

前提——电子化的高速公路

**First There Were
Electronic
Highways**

1939年，通用汽车公司还只是一家大胆创新、创意营销的年轻公司，就敢于在纽约皇后区的世界博览会上向世人展示了世界上第一台无人驾驶汽车原型。公众为之兴奋不已，然而，直到六十多年后，通用汽车的梦想才由谷歌功能强大的无人驾驶汽车实现。

无人驾驶汽车的发展迟缓了几十年，这是一种被称为“达·芬奇难题（Da Vinci problem）”的现象，而这个难题在今天依然困扰着发明家。达·芬奇难题指的是发明家设计了一种还不能应用的科技产品，之所以不能应用不是因为理论有问题，而是因为关键性的使能技术还未被实现。也就是说发明家的设想是合理的，但必备的技术还不具备。有个典型的事例。1493年，列奥纳多·达·芬奇（Leonardo Da Vinci）就设计出了直升机，但可惜的是达·芬奇的这种垂直起降飞行机器（他将其称为“螺旋桨”）并没有飞起来。它的设计原理是可行的，达·芬奇的设计图纸也没有问题。致命缺陷在于，为了能够起飞，达·芬奇的螺旋桨需要一个轻巧

而又牢靠的机体和强有力的动力源，而这些在当时都还没有被发明出来。现代的直升机也是直到20世纪才得以起飞起来的。

自从垂直起降飞行器被达·芬奇设计出来，历经几百年，直到有了现代化的内燃机和轻型铝结构机身的搭配，它才最终能够飞起来，无人驾驶汽车的发展也经历了类似的过程。早在1939年，通用汽车公司的市场部人士就设想出了第一代无人驾驶汽车，并在世界博览会上公之于众。博览会占地1200英亩，展现了众多令人兴奋不已的前沿技术，如电视机、电子路灯、荧光灯，以及新晋中产阶级家庭中必备的设备——自动洗衣机。通用公司大胆地展示了一个名为“未来世界”的作品，即一个自动化的高速公路，并预计到1960年时它将使驾驶出行可以“解放双手，解放双脚”。

通用汽车的“未来世界”以美国的一个典型城镇为背景，构建了未来现代化后的沙盘模型。参观者凭借移动座椅的便利，随着椅子移动欣赏这个迷你世界的动人景色。迷你的城市、农场、郊外，甚至还有一个小机场，接连从参观者的眼前滑过，而将一切这些建筑连接起来的是宛若绸带穿行于其间的“自动化高速公路”。在18分钟的浏览过程中，参观者还能听到一段解说录音，描述着这个“未来世界”的幸福生活。

“未来世界”的解说词为人们描绘了1960年的生活：普通人也可以享受便利的出行，所依托的就是在自动化高速公路上行驶的借助一套无线电控制系统管理的汽车（这些受无线电控制的汽车具体的工作原理被有意含糊过去了）。按照解说的设想，仅仅在据当时21年后的未来世界，人类驾驶员就都会变成休闲的乘客，这都要感谢通用汽车公司研发出的创新汽车。无线电控制的汽车会自行引导驶进驶出自动化的公路，既安全舒适又

经济便利，将人们从家里送到公司、机场，或者任何人们想去的地方。

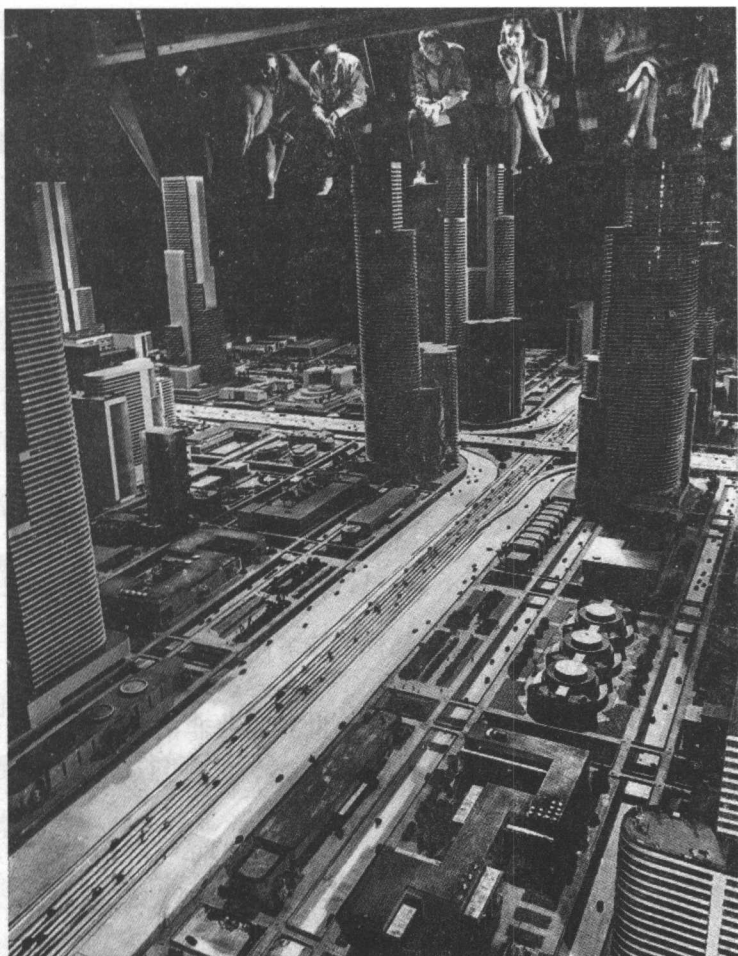


图 5.1 “我们现在看到的迷你模型是一个设想、一个艺术构思，但一旦它成为明天的现实，将会引领世界发生诸多变革”，纽约世界博览会，“未来世界：公路与地平线（Futurama: Highways & Horizons）”，1939年

来源：通用汽车公司

今天，我们很难设想就这样一个主要卖点是高速公路和迷你城镇的景观模型，能吸引数百万人的热切追捧，可是在1939年，一半以上美国民众都沉醉于通用汽车这个自动化高速公路的乌托邦设计中而不能自拔。通用汽车的“未来世界”是1939年世界博览会上最成功的作品之一，据事后统计吸引了数百万人的参观。展出之际，曾在一天里有2.8万人排队数小时等待参观，有时等候队伍还会绵延两公里。

通用汽车的这个人造迷你世界就像是个大剧院，犹如问世于同一年的电影《绿野仙踪》（*The Wonderful Wizard of Oz*）中的奥兹王国一样魔幻、有魅力。世界博览会上展现的其他高科技成就最终都走进了人们的日常生活，与之不同的是“未来世界”的自动化高速公路虽然精彩夺目，但也正如《绿野仙踪》中的飞天猴子一样不那么现实。对于这些“解放双手”的汽车的具体工作技术细节语焉不详，只是将这种引导汽车自动驾驶的能力含糊地归结为“无线电和电子学的某种巧妙结合”。

其实，只要知道通用汽车的“未来世界”并非由工程师构思设计，而是由舞美设计师出身、工业设计界的传奇人物诺尔曼·贝尔·盖迪斯（Norman Bel Geddes）设计出来的，对于最后的这个结局也就不足为怪了。贝尔·盖迪斯之所以为人所知，就在于他尤其擅长营造魔幻电影中的场景布置，对鸡尾酒调制器、台灯、收音机等日常物品进行未来化诠释。在那次博览会之后，贝尔·盖迪斯写了一本叫作《魔法高速公路》（*Magic Motorways*）的书，讲的是如何将这种精心设计的高速公路变成现实。他认为1939年展出的“未来世界”之所以能备受大众追捧，就在于人们渴望汽车带来的隐私性和出行便利，但又吝于为此付出的代价：

成千上万的人排队等着乘坐汽车，但也因此承受各种烦恼——每天忙着从一个地方赶往另一个地方，拥堵的路口、堵塞的窄道、危险的夜间驾驶、烦人的警笛、鸣叫的喇叭、闪烁的交通灯、误导人的路标以及令人气愤的交通规则。人们震惊于每天频发的交通事故及其造成的伤亡，所以渴望能有个明智的解决之道，让他们摆脱这种近乎于盲目自杀的混乱行为。

很多现代的观察家可能会对这些有关早期路况的负面描述有些惊讶，毕竟在那个年代汽车还是昂贵的奢侈品，仅仅是一小批富裕的人才买得起。然而，即使在1939年，这些交通问题就已经在上流社会中出现了。在城际公路网还未建立之前，在黄绿色的汽车尾气还未覆盖洛杉矶和北京这样的大城市之前，人们就已经发现汽车是一种不安全的、高代价的个人交通工具，会引发“盲目的自杀”。

贝尔·盖迪斯认为经济的繁荣、军备的完善、公众安全感的提升，均依赖于精心规划设计的公路网。除了带来实际利益外，贝尔·格迪斯还认为高效率的公路网系统对于营造国家的政治氛围，也是有着至关重要的作用。对他而言，第二次世界大战在即，法西斯独裁者在几个欧洲国家开始掌权，这样的时局下便利的个人出行是确保团结统一的根基。

贝尔·盖迪斯在书中热情地呼吁人们关注精心设计公路网的重要性，因为良好的路况是一个非标准化的自由社会的根基：

人们能够自由自在地在通畅的道路上行驶并享受体面的生活。人们能够自由自在地在各处旅行并借由开阔的视野感受到心灵的馈赠。生活在这样的美国所带来的内心的转变也远胜于外在的形式变化。如果城市居民可以感受乡村气息，东海岸的人能与西海岸的人相识，居住在山区的人能领悟海港和大海的味道，人们的眼界眼光得以拓展，寿命也随之延长，那么这些丰富而深刻的变化还会引领出另一种事物——我们称之为团结统一。这种团结统一不是自上而下的要求一致，不同于独裁专制下的压迫使然，而是基于自由和相互理解建立起来的内心团结。

就在贝尔·盖迪斯写下这些激动人心的话语之后不到一年，美国正式宣布参加第二次世界大战。通用汽车也随之将自己的精力从建造自动化高速公路上移出，转向为同盟国制造坦克、飞机以及其他武器。自动化高速公路项目在20世纪40年代几近中止，但事实证明，第二次世界大战无疑为后续几十年的发展提供了可供持续挖掘的巨大技术宝库。

战争是科学技术的比拼，政府在军用雷达、电脑以及激光上投入大量人力物力，为后续几项核心技术的发展奠定了基础：战事结束后很长的一段时间里，这些战时所用技术都被用于加速20世纪五六十年代电子化高速公路的发展，并于80年代助力第一代全自动化驾驶汽车的上路。

黄金时代

对于汽车和电子化高速公路的发展，20世纪五六十年代都是一个黄金时代。伴随着二战后的经济繁荣，成千上万的人购买了他们的第一辆车。根据美国的人口普查的数据显示，截至20世纪50年代末，大多数美国家庭都至少拥有一辆汽车。

为了鼓励汽车运输业的发展，1956年联邦政府通过了《联邦高速公路法案》（*Federal Highway Act of 1956*），开启了持续数十年的公路疯狂建设，并最终让美国各个城市、郊区以及乡村的格局面貌焕然一新。随着购车用户数量的急速上涨，数百万驾驶员奔向了个国家新建的各个道路和高速公路。随着数万英里的道路铺设，横跨东西海岸的州际公路系统的完善，美国的“汽车文化”就这样诞生了。

人们积极拥抱汽车给他们带来的低成本、便利的个人出行，让贝尔·盖迪斯自动化道路带来的团结一致的诗意描述成为现实。高效的高速公路和汽车真的为城市居民提供了领略田园风光的机会，住在山里的人们也可以亲身体验海港风光和海风吹拂。但没那么诗意盎然的是，车轮上的国家也开始在新出现的快餐店里采买食物——人们驾车穿过汉堡店柜台，买好食物后享受车中进餐的便利——汽车成了一种新型的起居室。到20世纪50年代末，全美国诞生了4000多家汽车影院，占据了全国电影屏幕总量的25%。

随着越来越多的人在车里饮食、娱乐，汽车公司也竞相制造功能强大并且空间宽敞的汽车——外观精心设计打磨，车内放有舒适的沙发。今天的汽车也许从技术上要比当时的领先许多，但是对于一些人而言，现在的汽车缺少早些年的设计灵魂。相比之下，20世纪50年代的汽车在设计上更追求美学的享受，它们那些体现曲线美的多元流线型车身均是效仿于喷气机和火箭的外观设计。

最近，在去拜访一位相隔数小时车程的亲友的路上，我们被路边一个古董车展的宣传所吸引，停下车去看了看。我们把租借来的黑色本田小汽车（经济实用，但终归还是有些紧凑）停好，朝着一个停车场走去，那里停放着几排20世纪五六十年代的古董车，它们全都经过打磨处理恢复了昔日的风采。老式的别克、雪佛兰、凯迪拉克威严地成排列队，这些都是石油禁运、大宗产品召回以及海外汽车制造商的竞争还未出现前的辉煌遗迹。老式的大型汽化器和泛着光的线圈管道被剖开展示，成群的人围在下面观察，但多半都是上了年纪的人。如同灭绝了的鸟类还在炫耀翅膀上羽毛的光鲜亮丽，这些闪闪发光的金属机体自信地展示着它们的光彩，亮红色、橘黄色、水晶蓝在土灰色的停车场地面映衬下显得格外耀眼。

当20世纪50年代很多美国人都在享受汽车带来的出行便利和私密旅行时，人们其实并不想承担驾驶的任务，自动化高速公路的设想在公众的观念里依然流行。50年代中期，电力公司合并后的广告宣传才让人们意识到电力消耗在飞速增长，而这在当时被视为经济繁荣和社会进步的代表性标志。有这样一则平面广告生动地展现了人们对于电动汽车的设想：一家人面对面坐在车后座上，围着一个桌子玩多米诺骨牌游戏。搭配的文字说明如下：

有一天，你的汽车将行驶在电子化超级公路上，它的加速和转向全都自动地由内嵌在道路中的电子设备控制。公路将会变得安全可靠，这都要归功于电子化！没有拥堵，没有事故，没有疲劳驾驶。

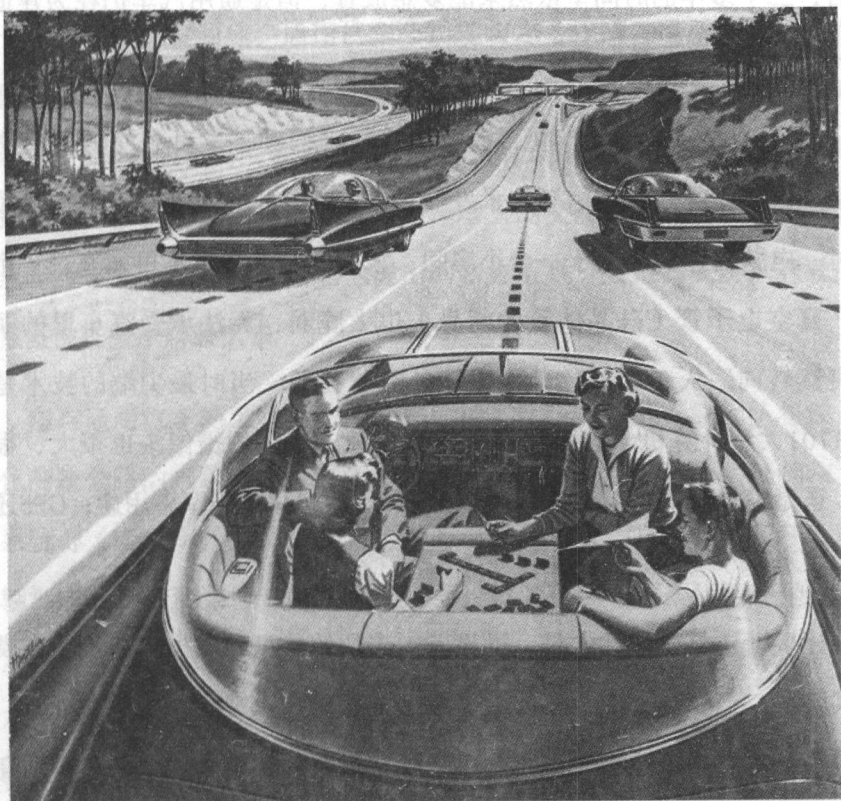


图5.2 “电子成为驾驶员”，美国电气照明与动力公司的广告创意“未来的无人驾驶汽车”，《星期六晚报》（*Saturday Evening Post*），20世纪50年代
来源：《埃弗雷特收藏集》（*The Everett Collection*）

1958年，迪士尼公司在其滚动播放的一个电视节目中插入了这样一段情节，彰显迪士尼乐园建造了“美国的魔法公路”。动画片节目的最后几分钟

里，可以看到一辆家庭用车正自动地在车库里清洗并充电。然后，全家人钻进车里，爸爸拽出几个杠杆来编辑汽车的目的地，汽车就自己开动了。行程中，爸爸在与其他人进行电话会议，而妈妈和孩子们则在后排休息。

继通用汽车在世界博览会上成功地展出“未来世界”让人们神魂颠倒之后，二十多年的时间里依然未能梦想成真，但是通用汽车仍在为建造自动化的高速公路而努力。

让公路来引导汽车行驶，而非汽车自动驾驶，这种想法看似离奇，但是如果你结合当时的信息技术发展状态来看，也就会觉得合情合理了。那时，电脑的体积相当于一个小房间，太大太重，根本不可能塞进一个车里。真空电子管（硅晶体管的前身）也太脆弱，无法承受汽车里的颠簸和突然转向。人工智能软件更是还处在萌芽期。当时最尖端的技术是教IBM701大型计算机学会西洋跳棋或扑克牌。当时的摄像头也很大、很脆弱，还是模拟制式（Analog）的；激光和雷达仍只限于军事用途；GPS还未发明出来。

让事情更具挑战性的是，世界上大多数的信息和数据也还停留在模拟制式阶段。当时，文本数据和数值数据要么是人工观察并记录在纸上，要么是借助工具费事地测量观察然后记录在纸上的表格里的，仅有个别高级的地方，会将表格以电子表格形式存储在大型计算机上。没有数据，没有详尽的电子地图，没有快速的电脑或准确便捷的传感器，将还处于原始阶段的人工智能插入到两吨重的钢铁机器——也就是汽车里，实在是不可能完成的任务。既然计算技术当时十分不成熟无法承担驾驶任务的监控，通用汽车的工程师就另辟蹊径，想出了其他解决方案：彩色电视技术。

通用汽车的电子化高速公路

美国无线电公司（Radio Corporation of America, RCA）是20世纪50年代电子工业创新的发源地，通用汽车公司遂与之合作研发电子化高速公路。美国无线电公司雇用了当时著名的发明家弗拉基米尔·兹沃雷金（Vladimir Zworykin）——阴极射线管及自动化解决方案的开拓者，由其研发信号灯的管控系统。兹沃雷金很早就意识到高速增大的交通密度和公路上飞速奔驰的汽车会带来负面影响，因此他提出的新方法是“要将驾驶员从驾驶的机械重复性任务中解脱出来”。

在与通用汽车的工程师团队合作之后，兹沃雷金的专家团队将设计自动化高速公路的难题拆解为三个子问题。正如兹沃雷金所言：“任何汽车控制系统都（必须）得知道公路上每辆车的具体位置信息。这就意味着需要车辆拥有某种侦测手段。第二个必备条件就是每辆车都必须清楚地知道自己及周边汽车的位置。通常车辆是尾随前进的，因此这就需要在车辆之间或车辆与道路之间建立某种沟通手段。第三步就是要汽车的自动控制系统能针对接收到的信息作出回应。”

经过几年的研究，两组团队拼凑出来一个他们称之为电子化高速公路的创意解决方案（尽管以今天的标准而言，还是略显粗糙）。他们结合了无线电技术、电子电路以及基于历史悠久的电磁学理论而创建的逻辑门电路。1958年，他们完成了电子化高速公路的一次最全面的展示，在内布拉斯加州林肯市郊区一条400英尺长、专门改造过的高速公路上，装有俏丽垂

直尾翼和双摄前大灯的两辆1958年款雪佛兰参与了测试。

凭借内布拉斯加州道路管理局的大力支持，通用汽车和美国无线电公司的研究团队组装出一套原始的车辆侦测与引导系统，可以实现在两个主轴方向上对汽车移动的控制：横向维度，将车辆稳定在车道界线内部；纵向维度，将车辆与前后其他车辆保持一定的安全距离。两年后，利用同样的方式，他们又在新泽西州建立了另一条电子化高速公路的测试跑道，但这次是为通用汽车专门定制的，最终成功实现了车辆的自动启动、加速、转向与停止，全程没有人工直接参与。



图 5.3 自动化公路系统测试，20世纪50年代

来源：美国无线电公司，大卫·沙诺夫图书馆（David Sarnoff Library）

1960年，新泽西州普林斯顿市政府发表了一篇激动人心的新闻稿，对促成电子化高速公路的技术进行了解读说明。

车辆侦测系统可以实现前后车辆自动维持安全距离的功能，并在道路前方出现障碍物或停泊车辆时自动停止车辆向前行驶。引导系统则需在行车道的中心铺设一条连续延长的电缆。

汽车前端装有两个用于距离引导的电磁线圈，负责拾取电缆中的信号电流。只要汽车的行进方向偏离车道中心线，一侧的信号强度就会强于另一侧，并产生“差异信号”提示，这个提示既然可以通过闪光或声音提醒驾驶员，也可以自动地引导方向盘调整。

车辆侦测系统是各种复杂电子设备的综合。这个系统必须实现基础设施之间的信息通信，这涉及晶体管、无线电发射器以及灯光。为了创建这一侦测系统，美国无线电公司的工程师首先在道路中铺设一系列矩形电线回路，这些矩形回路的长度比汽车车长略短，一个个矩形回路之间首尾相接，覆盖整条测试道路。每当车辆行驶经过一个矩形回路，就会向埋在道路中的晶体管侦测设备发送一个特殊的信号。

当车辆快速驶过这一系列矩形回路时，反馈信号就会流入由所有侦测器组成的一个网络里。这一系列连续的动作信号会点亮道路边缘的一串照明灯，形成一种电子化“飞行尾翼”的感觉，进而起到警示附近其他车辆的效果。从理论上讲，人类驾驶员能够清楚地看到无人驾驶车辆的位置就是缘于两侧路灯的照明。其实，也并不需要人去注意，这些信号就会以无线电的形式传播给附近的控制塔，然后由控制塔自动地以无线电指示的方

式传递给附近的前后车辆，提醒它们通过刹车或加速来调整与这辆车之间的距离。

构建控制车辆在车道中加速或减速的引导系统，需要涉及略微复杂的工程学，并进行富有创造性的组合。为了模拟人在掌握方向盘时的关注焦点和神经反射过程，通用汽车和美国无线电公司的联合工程师团队将电子配件与电磁现象进行巧妙组合。而为了收集汽车在车道上的横向位置这一关键信息，联合工程师团队利用了现代电子技术，以及电流与磁场之间早已存在却未被深度理解应用的关联效应。

一个多世纪前，科学家就已经知道电流可以产生磁场。反之，变动的磁场可以在附近的任何导电材料中引发电流。导电材料距离磁场越近，产生的电流越强。当物理学家依然在努力探索为什么一个通电的线路会辐射出一个磁场的具体原理时，科学家和工程师早已学会了如何测量并熟练地应用磁力和电力这两种原始力量之间的关联效应。为了对线路中流动电流产生的磁场范围和强度进行量化，科学家和工程师使用了众所周知的安培定律。反过来，为了对磁场会引发导电材料产生多强的电流进行量化，专家们应用的是法拉第定律。不足为奇的是，安培定律和法拉第定律只是证实了大多数人直观感觉到的现象。导电路径中传导的电流越强，引发的磁场也越强。同样的道理，在导电材料附近移动的磁场越强，导电材料中被激发的电流也越强。

现在对于电流和磁场之间的关联效应应用最为人熟知的一个实际产品就是隐形的狗围栏。住在郊区的人们为了拴住他们的狗，会在他们的房产周围埋设一条通电线路。埋下的线路中会通入一股振荡的电流，以便辐射

出一个振荡的磁场。在一副特制的狗项圈中嵌入一个导电金属传感器后，它就会与振荡磁场发生相互作用。一旦家犬套上这个特制的项圈，当狗靠近房子的外围（和埋设的线缆）时，狗项圈内的传感器就会激活并实施一次轻微电击，最终让狗学会待在自家的院子里。

与被人喜爱却又经常四处走动的家犬类似，不受任何显著围栏阻挡的无人驾驶汽车是最自在的，但出于安全考虑仍需限定其在特定的地理区域内活动。也就是说，行驶在通用汽车和美国无线电公司联合研发的电子化高速公路上的无人驾驶汽车，只有限制在车道上的安全范围内，才会取得最理想的表现。早期自动驾驶的汽车所使用的引导系统就是利用了隐形狗围栏的原理。不同于埋设在房子周围的通电线路，美国无线电公司和通用汽车公司在他们设计的电子化高速公路测试跑道的车道中间埋设了通电线路。下一步类似于在狗项圈内嵌入传感器，工程师为每辆汽车装备了两个金属“传感线圈”，等距地分置于汽车的两侧。与每个传感线圈匹配的是是一套测量设备，用于对其中通过的电流强度进行量化。

请设想如下的场景，装有两个传感线圈和测量装置的一辆汽车缓慢地行驶在测试道路上，当汽车从道路上方驶过时，地下电缆会产生磁场，而这个磁场又会引发车载传感线圈产生电流。如果车辆正确地行驶在道路中央，两个传感线圈中产生的电流将会大致相当。然而，如果汽车危险地偏向了一侧，这侧的传感线圈就会产生更强的电流，对应的传感器也会记录下相较于另一侧较高的数值。接受到较强信号的传感器就会向汽车的方向盘操控系统发送指示，要求车辆轻微转向，直到两侧的传感器测量数值再次平衡。

这一过程就被称为“反馈控制（Feedback Control）”，在当时被视为一项十分先进的技术。借助汽车传感线圈中的传感器记录数据，并监测测量数据的稳定与否，这套技术的组合构成了一个虽然略显粗糙但行之有效的自动操作系统。早在计算机机器视觉发明以前，通用汽车和美国无线电公司共同搭建的这套方向盘控制系统就可以自动调整车辆在道路上行驶时的横向位置，其准确性和反应速度丝毫不逊于一个注意力集中的人类驾驶员。

当时，对于通用汽车和美国无线电公司的这套粗糙的原型系统，人们的乐观情绪高涨。普林斯顿市政府甚至发表了一篇通讯稿，热情豪迈地认为在未来的某一天，这一套电子化高速公路的发明将使得“未来的出行者可以在周末出游时坐在车中打桥牌或睡一觉”。然而除去公众的热情反馈，通用汽车的电子化高速公路方案并没有真正实施起来。

不过，通用汽车和美国无线电公司的长期合作还是带来了一些其他贡献。在今天，电子化高速公路方案中车辆侦测系统的核心理念仍广泛应用于反馈式信号灯的制造。自动化信号灯之所以能调控车流，关键就在于它能侦测出路口是否有汽车出现，或者有车辆等待左转。当侦测到车辆时，由埋地线路和传感器组成的一套回路系统就会发送电子信号，让信号灯由红变绿。

也许是受到先前成功的鼓舞，通用汽车公司在1964年的纽约世界博览会上再次作出了无人驾驶汽车的尝试，而那里正是25年前1939年“未来世界”大放光彩的地方。通用汽车在广告中这样宣传他们的概念车“火鸟”：“某一天，一家人驾车行驶在超级公路上时，可以将车辆的控制权交给一套自动程序化的引导系统，人们随后就能尽情享受旅程的舒适和绝对

的安全，并以今天高速公路上汽车的两倍速度前往目的地。”尽管火鸟有着吸引人的柔美线条和单峰垂直尾翼，它还是成为了通用汽车在无人驾驶汽车领域投入几十年后的绝唱。

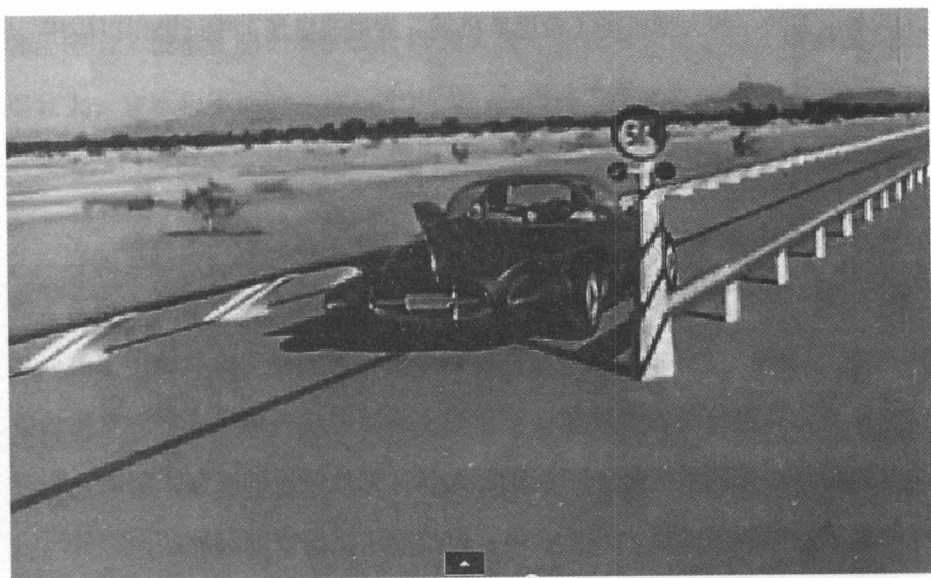


图5.4 以“涡轮动力”驱动的通用公司“火鸟”概念车驶入一条自动导航车道，来自1956的通用汽车展览会

来源：通用汽车公司

纵观20世纪六七十年代，利用通用汽车和美国无线电公司的基础系统——电缆、金属传感器和磁场感应器的组合，其他研究者继续针对自动化高速公路做了些改进升级。20世纪60年代，美国俄亥俄州成为了汽车工程领域的前沿研发中心，在当时尤其以自动化车辆引导和控制的研究为人所熟知。同一时期，英国交通道路研究实验室（Transport and Road Research Laboratory）测试了一辆雪铁龙DS款改造的无人驾驶汽车，区别在于只是将引导线路嵌在测试跑道的表面。

直到20世纪70年代末，大多数无人驾驶汽车的研究也还没有将重心放在小型汽车上。后来，研究者将精力投入到无人驾驶汽车的另一细分领域中——自动化大众运输工具（automated peoples movers, APMs），而这一套方案在今天的一些大型机场（希思罗机场、肯尼迪机场）中已经得以应用。

电子化高速公路的时代也随之终结。

自动化高速公路的衰亡

自动化高速公路的美梦最终破灭的主要原因之一就是成本。安装必备的电缆和路边控制系统是一项耗资巨大却又见效缓慢的工程，装配一条短小的测试跑道所需的成本还算合理，但是对于美国或欧洲那些横跨各州的浩大公路网系统而言，方案就显得不切实际了。即使在20世纪60年代，政府划拨了大量的公路建设预算，但考虑到埋设线缆、晶体管以及其他电子设备组成的这一套脆弱的基础设施，想要装配并维护数万英里州际公路的成本也太过高昂了。

在1969年发表的一篇文章中，俄亥俄州的研究者罗伯特·芬顿（Robert Fenton）和卡尔·奥尔森（Carl Olson）就提出“每一英里基于计算机和传感器搭建的高速公路，需要投入的总成本在2万到20万美元之间。”电子高速公路不仅成本太高，而且当时可用的电子技术和计算机技术都还太粗糙。尽管芬顿和奥尔森都是自动化高速公路的拥趸和专家，但是他们也不得不承认即使经过了数年的紧密研究和快速发展，“目前仍无法完全罗列出实

现该方案所需的系统元件，因为还缺少关键性的技术知识”。

不考虑成本和技术因素，致使身处黄金时代的自动化高速公路项目最后走向终结的另一个原因，是汽车行业失去了最初的理想主义。汽车行业曾经吸引了很多大胆创新的设计师和大量专业技术人员，但随着自身的发展壮大，它不得不面对这样一个事实——汽车不再是新鲜事物了。汽车和高速公路已经成为每天生活的实用工具。

关心消费者的律师拉尔夫·纳德（Ralph Nadar）于1965年发表了成名作《任何速度都不安全》（*Unsafe at Any Speed*），揭露了大型汽车公司的一些伪劣工程产物，细致地讲述了雪佛兰科威尔汽车的安全问题，以及汽车行业是如何关注款式设计和利润空间而忽视驾驶员安全的。这本书成了当时的畅销书，并且促成了《1966年国家交通和汽车安全法》（*National Traffic and Motor Vehicle Safety Act of 1966*）的颁布，以及安全带法规在49个州（唯一的反对者是新罕布什尔州）的实施。引发更多争议的是几年后的调查，美国交通部针对几种小轿车做了一系列的测试后得出结论——尽管纳德做出了指控，但是相比“同时期的进口车和国产车”，1960-1963年产生的雪佛兰科威尔系列车的安全性还算是较好的。

虽然科威尔系列车得以免罪，但是已经无法抹去公众对通用汽车、福特汽车和克莱斯勒汽车的疑虑。随着高速公路上的汽车数量持续增长，消费者安全运动继续发酵。到1970年时，已经有数百万的城市居民迁移到了郊区居住。结果就是，每天通勤去上班的人数翻倍，从1950年的3600万增长到1970年的7400万。这既是财富增长的象征也是诅咒的种子，美国州际高速公路网系统已经演变成错综复杂的基础运输设施，并面临着一个新的

挑战：交通拥挤。

1973年，随着几个石油出产国为了对抗美国给予以色列的支持，将每桶石油的出口价格上调了70%。廉价石油的时代突然结束了，美国汽车工业曾经的辉煌便一去不复返。大型汽车公司逐渐放弃了曾经对“解放双手、解放双脚”的自动化高速公路的大胆幻想，转而关注更加切实的挑战难题，如制定汽车安全标准，提升燃油效率，减少尾气排放。

从20世纪70年代到80年代，无人驾驶汽车仍还只停留在概念阶段，并没有成为现实。随着汽车工业焦点的转变，流线型的设计成为了继燃油效率之外的第二个焦点，利润诉求击败了人们的美妙设想。尽管大量有才华的工程师继续为此努力，但20世纪余下的几十年里，受限于当时信息沟通技术的不成熟，无人驾驶汽车的发展依然举步维艰，始终无法突破达·芬奇难题。

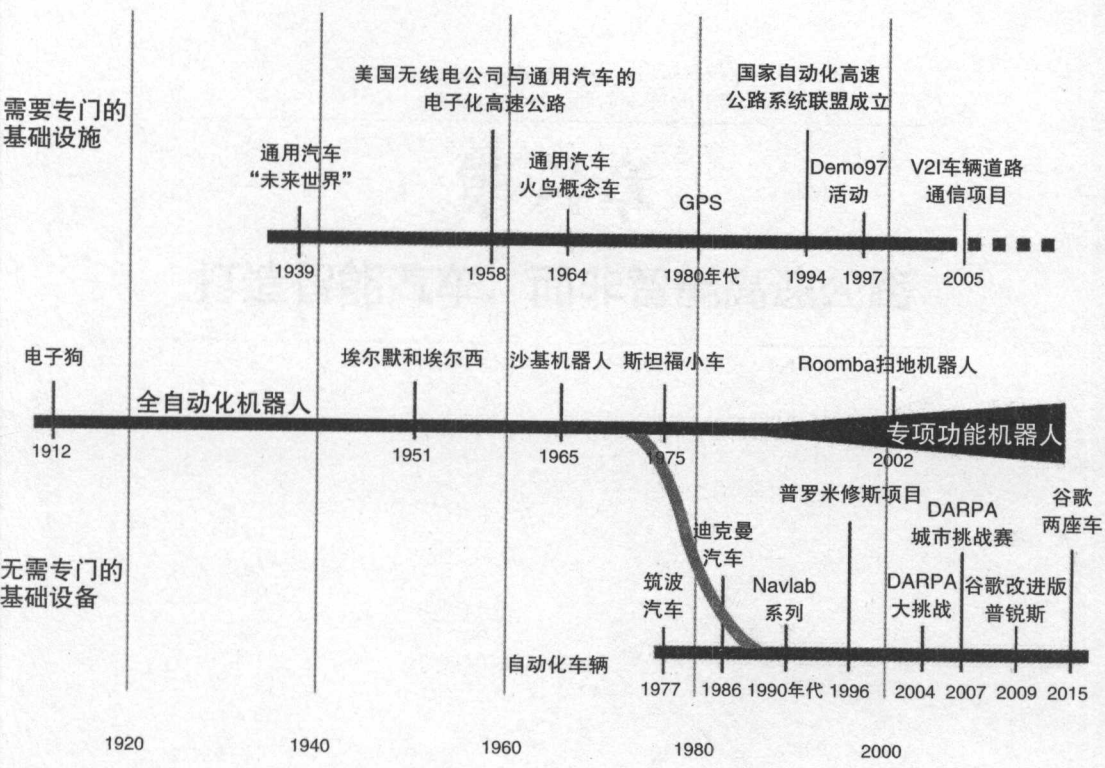


图 5.5 自动驾驶汽车的发展里程碑

第六章

打造智能汽车，而非智能高速公路

**Build Smart Cars,
not Smart
Highways**

在电子高速公路的黄金时代过去了近三十年后，一场智能汽车会议在华盛顿某酒店的无窗大会议室内召开。负责国家航空、铁路、高速公路、汽车政策的美国交通运输部组织众位行家前往参会。我是250名参会者之一，结果我们的时间却都耗费在聆听西装革履的专家们轮番上台的夸夸其谈上。

我来参会的目的是为了加深对无人驾驶汽车的了解。然而，接连好几位演讲者却用幻灯片给大家展示各种晦涩难懂的缩合术语：RSU、FCC、TMC，还有V2I，纯粹在拖延时间。失望至极的我在早会茶歇时，留意到一小群人在咖啡茶点桌旁小聚，于是我向他们走去。他们看起来挺靠谱，似乎很清楚自己在做什么。

自我介绍过后，我在交谈中获悉，他们当中有汽车行业顾问、汽车工程专业博士后，还有一位运输技术教授。我不解地请教他们，为何没有一则演讲提及计算机视觉或是深度学习等相关话题。事实上，当时已经开始试验的谷歌的自动驾驶汽车也未曾被提及，一次都没有。

下水道盖子般大小的吊灯在我们头顶散发出炽热的光。灯下是那群专家学者的直眉瞪眼，仿佛在围观一个脖子上长了两个脑袋的怪物。长长的沉寂过后，那位汽车行业顾问终于开了金口，解释道：“噢，无人驾驶汽车，我想明天下午就会有单独配置的‘轨道’了吧。”在我还没明白过来的时候，他们却话锋一转，继续之前的讨论。我趁机离开，心有戚戚。我不如他们懂行，却坚持自己的想法。

多亏了这次美国交通运输部会议的经历，让我凭直觉意识到“智能汽车”这个词对于联邦交通运输机构内部和大学机器人专家有着不同的含义。从国家政策制定者的角度出发，对于“智能交通运输系统”制定的政策并不一定包含机器人。事实上，在过去几十年中，美国联邦交通运输机构筹资成立的大部分研究所都致力于发展“车对车通信系统（vehicle-to-vehicle, V2V）”与“车与基础设施（vehicle-to-infrastructure, V2I）”。从广义上来说，“车对外界的信息交换（vehicle-to-everything, V2X）”越来越趋向指V2V与V2I计划。

V2X技术

美国联邦交通运输机构关于V2X的倡议，目标在于创建无线交通基础设施网，连接每辆汽车。它们通过路边传输器共享数据，从而减少事故，挽救生命。V2X研究利用的是由联邦运输委员会（FCC）监管的专用短程通信技术网络（DSRC）的部分储备带宽。负责V2X计划的机构是美国交通运输部的一个主管部门——美国高速公路安全协会（NHTSA）。

通常情况下，当我觉得研究会议无聊时，会迅速抽身止损，看看能否尽早乘坐航班回家。然而这一次，不确定的因素太多了，我必须留下继续学习。

午饭休息过后，我返回会议室听了好几场充斥着各种缩略术语的演讲。一天下来，我发现每年高达上百万美元的税收被用作数据交换研究的资金。未来五年内，美国交通运输部计划再投资1亿美元，用于车联网环境（connected vehicle environments）的设计与应用项目试点。可问题在于，他们的研究类型错了。

在美国，无人驾驶汽车的命运掌控在美国交通运输部，及其两个关键的分支机构——联邦公路管理局和美国高速公路安全协会手上。根据1966年出台的美国交通运输部的相关法规，其下属部门有权通过联邦命令，强制50个州政府和汽车公司同时遵守。至少从理论上来说，美国交通运输部的职员也可以构想出推动无人驾驶汽车发展的法规。毕竟，历史上该机构曾不顾社会大众的反对，坚持推行座椅安全带，使之最后成为汽车安全的必需配备品。

美国交通运输部的会议结束后，我问自己：“为什么不减少投入在汽车连通性方面的资金，增加对于机器人技术的研究经费呢？”美国交通运输部2014年的预算高达770亿美元。其中410亿美元拨给联邦公路管理局，8亿美元投入到国家公路交通安全管理局管理的交通安全项目（含汽车、街道、桥梁、隧道和高速公路）。仅用从这笔巨款中拨出一小笔资金用来支持无人驾驶汽车研究，就可能实现许多技术进步。

从华盛顿开会回来，我增长了一些知识，但更深入研究V2X背后的技术

后，我却愈发感到悲哀。V2X倡议的潜在目标是安全性，美国交通运输部官网上相关表述如下：

通过匿名交换定位、速度和地点等相关数据，V2V技术能让车辆感知周围360度范围内其他车辆的定位以及带来的风险，计算风险，并对司机提出建议或发出警告，预先采取措施避免或减少车祸的发生。

应用了V2X技术的汽车和基础设施会配备专用短程通信技术网络的收发器，在300米范围（约1000英尺）内传输数据。连通的车辆和基础设施能在75兆赫的划定区域内实现交互，频率限定在5.850~5.925千兆赫之间。一辆V2X技术的汽车每秒可传输10次，数据内容包括车速、位置、车型、前行方向、制动器状态等等。常见的警示则包括对恶劣天气和前方路障的提醒。而V2I技术能捕捉交通车流量的数据，及时调整周边的交通信号灯，提醒其他司机选择其他道路等。相关技术还包括专用短程通信技术网络的道路装置、其周围的光纤电缆和V2I交通信号灯。

客观地说，V2X技术的发展并不可怕。事实上，该项技术背后理论前景诱人。美国高速公路安全管理局的一项通过模拟潜在车祸场景的研究发现，如果所有的车辆都具备V2X性能，当司机（人类）在驾驶中遇到盲点和易忽略的障碍，系统会发出“哔哔”警报声。一年下来，这项技术预计能有效避免400万起车祸，占全年车祸数的79%。这可减少一大笔费用损失，并能挽救许多性命。

让我失望的是美国投入到“智能汽车”的资金非常有限。为什么短距离无线网络的路况数据交换——这项常见的、实用的技术应用——会成为“尊贵的”美国交通运输部“应用信息技术提升安全驾驶”政策的关键点呢？专注V2X技术，排斥其他智能技术，联邦交通部当局依然沉迷于追求某种技术上的范式（Paradigm）。从概念上看，该技术仅仅比20世纪50年代使用的无线电监控高速公路的初始水平提升了十分有限的水平。

在过去，美国交通运输部并不保守。事实上就在十几年前，该部门还曾大力推进智能汽车的发展。但问题在于这群思想超前的交通运输专家，犯了与当时的通用汽车工程师同样的错误：他们对于无人驾驶汽车的愿景设想大大超前于现实，导致缺乏完备的驾驶技术来支撑。

智能交通系统的历史

早在20世纪80年代，信息技术重塑了整个工业格局。1986年，为了解决美国日益增长的交通和汽车尾气问题，加州交通局联手加州大学，共同探索信息通信技术在汽车与高速公路的应用，帮助人们提升驾驶效率。这次合作最终促成了一个名为“先进技术与高速公路（Program on Advanced Technology and Highways, PATH）”的全国项目。很快地，加州交通局与该项目人员意识到没有联邦的支持，没有大型汽车企业的广泛参与，他们很可能徒劳无功。

然而PATH的利益相关者没有气馁，主持了好几个专题会议，并持续

推广。由于不堪忍受美国史上最糟糕的交通堵塞和城市空气的煎熬，1988年，来自联邦政府、多个州的政府机构、工业界和多所高校组成一个特别工作小组，名为“移动2000（Mobility 2000）”。该小组游说美国交通运输局尽快成立正式的联邦项目办公室，出台法律规章，促进先进技术应用，提升全国高速公路与道路的安全性及效率。到了20世纪90年代初期，为了促进智能车辆和公路系统（intelligent vehicle and highway systems, IVHS）的发展，美国交通运输局终于成立了正式的联邦项目办公室。

为了适应联邦项目的最新发展，后来更名为智能交通系统（intelligent transportation system, ITS）的IVHS项目，在当时可谓一项大刀阔斧的举措。它是一个成熟完备的联邦项目，涵盖所有自动地面交通模式：自动交通管理系统、驾驶员信息系统、商务用车和公共交通。同年，国会顺势而为，通过了一项重要的新交通法案——《综合地面交通效率法案》（*Intermodal Surface Transportation Efficiency Act*）。

交通运输部部长向公众宣布了该法案：“大力发展自动化高速公路，制定汽车标准。未来在此基础上推动自动化汽车高速公路建设……该项目的目标是到1997年之前，完成第一条全自动化公路建设或自动化测试车道并投入使用。”美国交通运输部把执行该远大目标的重任托付给了联邦公路管理局。于是，联邦公路管理局把该项目命名为“自动高速公路系统（Automated Highway System Program）”，分为三个阶段来实现：分析、系统定义和运行评估。

更为理想的是，国会批准了6.6亿美元资金用于未来6年间自动高速公路建模的研究、开发与操作测试。随着这些新投入的财物到位，自动汽车与

高速公路系统的发展仿佛必将取得激动人心的“车辆、道路、计算机”创新融合成果。然而，不幸的是，这种资金与法律的结合以失败告终，并没能让自动驾驶汽车的发展更进一步。

在美国交通运输部内部，尝试这种官方的项目既是祝福，亦是诅咒。一方面，交通运输部坐拥数十亿美元的财力，雇员多达55,000人。从理论上说，至少能提供良好的政策与资金扶持，推动无人驾驶汽车技术的进步；另一方面，把一个发展迅猛的科技密集型项目交给负责各种交通方式（航空、高速公路、公共交通和汽车）的庞大的联邦政府机构去管理，难免会带来风险——无人驾驶汽车技术会在这群“好心办坏事”的官僚手中渐渐衰亡。

在自动高速公路系统项目初期，人们普遍抱有乐观的态度。相关文件显示，在1993年就有部分美国交通运输局的雇员极其看好自动驾驶的前景。当时的人评价道：

这个高性能的高速公路系统，被视为下一个地面交通运输的重要发展阶段。人们期待它成为美国下个世纪初主要的交通工作重点，就像本世纪后半叶，州际公路系统计划成为大家的关注对象一样。

对比之下，今天的V2X项目就像在失落的自动高速公路系统文明中残存的遗迹。

到了1994年，此项目的资金已经全部到位。联邦公路管理局此时需要

交通专家团队参与进来，逐步设计、建造出自动高速公路的运行模型。这一年年末，美国交通运输局组建了拥有120名成员的国家自动化公路系统联盟（National Automated Highway System Consortium），来落实自动化高速公路系统的发展战略与最终切实可行的模型。发展至此，繁冗复杂的联邦项目管理体系终于形成，其中的各种机构、工作小组和团体数量之多，足以媲美以“复杂”著称的欧洲皇室家族族谱。

该联盟由九个主要机构组成，分别来自不同的联邦机构、大型汽车企业、电子协会、汽车运输部门、各高校、国家和地方政府、交通运输部门和咨询公司。从1994年到1997年，联盟循序渐进地尝试调和各成员的分歧，形成统一的组织议程，从而用一致的策略来管理高速公路系统。令人遗憾的是，利益各方非但没有团结起来成为目标一致、高效的团队，联盟本身却因为成员间目标不一、利益各异，不堪重负。

阵容庞大的联盟终于要开始吃力地往前挪动。分析、系统定义和运行评估，为了给项目的这三个阶段筹资，美国交通运输局分别与众位联盟成员签署了联邦合同。诱人的巨额联邦资金开始流动。高高在上的联盟成员排起队来，迫不及待要把资金收入囊中，另一边还有各大汽车技术巨头在等着：霍尼韦尔（Honeywell）、通用汽车（GM）、福特汽车公司（Ford）、PATH、德科电子（Delco Electronics）、卡内基·梅隆大学等等。

经过三年的艰辛付出与汽车、交通行家的努力，美国交通运输局终于成功定义和创造出“全自动化汽车-高速公路系统（Fully Automated Vehicle-Highway System）”，吸引了数千万的研究经费。为了展示其在过去三年内的研究成果，联盟在加州圣地亚哥北部举行了一场名为“Demo 97”的活

动，并邀请国会官员、政客和企业高层试驾原型车辆。

在“Demo 97”举办之前，群众的热情已经十分高涨。有篇文章满怀激情地写道：

未来的高速公路能让司机轻松地在开车时打电话，传真文件或阅读小说。自动高速公路系统把握着驾驶方向、刹车、油门踏板，让人们享受“解放双手，解放双脚”的轻松驾驶……国家自动高速公路系统联盟为我们描绘了这一美好的前景，也让交通安全与高速公路效率提升成为现实。

出席活动的可谓大腕云集，汽车工程师也极尽全力，大众更是反响热烈。在圣地亚哥蔚蓝无暇的天空下，“Demo 97”活动在15号洲际公路延长段举行。七天下来，观众看到了各式各样的汽车展示。每一种展示都以多种方式呈现出自动驾驶系统是如何实现车道保持与纵向控制（车与车之间适当的前后距离）。目的在于告诉人们，尽管没有人手控制方向盘、手动挡，没有用脚踩刹车和油门，自动驾驶车辆也能自主掌握方向，加速和刹车。

观众发出阵阵惊叹。道路磁铁系统的应用让人回想起通用公司和美国无线电公司几十年前做的实验。加州大学PATH项目组展示了由8辆1997年版的别克LeSabre轿车组成的车队，车子紧凑地排成一列纵队前进（车辆结队），这样驾驶可以有效节省汽油。本田公司展示了两款样车，配备了先进的敏捷操控辅助系统，能够在人和车之间转换驾驶方向（向前或向后），启用传感器自动变换车道，灵活避开路障。丰田公司给大家展示了

激光传感系统怎样工作，在遇到路障、盲点和变换车道时，系统会发出警示提醒司机。

一切都按计划顺利进行，车辆的表现无可指摘。众位高官政要，例如美国参议员和企业高层们放心地驾驶着展示的车辆。媒体对此充分肯定并大力宣传。诚然，“Demo 97”是美国交通运输部在发展迅猛的高速公路技术上最后的投资。但这次活动依旧有主要有两个问题：当时的无人导航驾驶技术还不够成熟，不足以保障安全；展示中的自动高速公路系统仍然需要配置昂贵的专门基础设施。

这次活动不经意间葬送了自动高速公路系统的前途。“Demo 97”之后，人们认为自动驾驶的想法不切实际。过了些时候，公众开始发现联盟的成果实际上几乎没有得出实际的结论和技术突破。于是美国交通运输局不再为自动化高速公路筹集科研资金，他们当时的想法在今天依然得到认同，即把司机辅助系统的科研技术当作扶持重点，而不是研究如何取代司机。但令人失望的是，美国国家交通运输局的“大胆”尝试依旧停留在V2X研究。

关于V2X的反思

联邦交通运输局拥有资源和立法权，能对无人驾驶汽车给予支持，这样每年能挽救成百上千个生命。2013年，美国高速公路安全协会朝全自动化的道路迈出了蹒跚的一步，并对外发布了详细报告。报告中列出了自动驾驶发展的几个理论阶段。这是一个良好的开端，接下来长路漫漫仍需努力。

尽管事实表明，在过去的十年，人工智能的发射器在软件和硬件层面都取得重大发展，这让无人驾驶汽车技术日趋成熟，有望实现，但高速公路安全协会仍然继续坚持发展V2X技术。事实上，就在2014年，协会郑重考虑通过政治管理手段，要求所有新上牌的轿车、卡车都配备V2X系统。以上这则规定，本意是好的，可是对提升道路安全没有什么实际用处。

让我们来详细解释在无人驾驶汽车中安装V2X系统有何利弊。V2X的好处就是安全性，至少单从理论上就是如此。预防交通事故是目标之一；另一个目标是提升交通效率，从而减少碳排放与尾气污染。

在理想情况下，V2X是这样工作的：当V2X汽车以特定的速度进入某条道路的交叉路口时，系统会发出通知。当它发出警示的时候，附近的其他司机就知道有车子驶近了，于是他们会刹车或减速。

从理论上说，车辆与基础设施之间的无线通信交互是非常实用的。如果红绿灯连接上网络数据，可以即时通知司机选择相应的道路；当司机听从系统提示，他们就可以据此作出加速或变换车道的决定；假如遇上恶劣天气，道路节点就会根据道路状况的信息发出广播，提醒司机注意路面打滑；桥梁、围墙、路障、路缘石都被清楚标记，提醒司机注意它们的具体位置，避免发生碰撞。

以上都是V2X的潜在优势，但现在我们来看弊端。

机器人专家会反驳：V2X导航的1亿美元的专款，本来可以用来研发更具创新意义、更智能的自动驾驶汽车。车联网的想法不仅技术不成熟，而且观念过时。在实际应用中，还有不少障碍阻止其在现实世界发挥应有的作用。

在我们看来，美国交通运输部现行的政策最大的缺陷是V2X只能在以下两种情况下提升道路安全和交通流量：汽车配置全自动化（指无人驾驶）；大部分汽车都在道路上行驶，而且所有的道路都安装了V2X设施。但眼下，我们无法达到这些要求。事实上，美国交通运输部把V2X当作汽车自动化的必经阶段，不仅实在浪费钱，而且可能恰恰妨碍了汽车全自动化的发展。

来说说钱的事儿吧。美国财政部关于V2X的报告显示，每安装一个用于想V2X提供数据的V2I道路装置就要花费高达51,650美元，这其中，连接道路设备与交通管理中心的光纤安装费用中占了很大比例。高昂的成本让美国的公共道路、隧道和桥梁都陷入了可悲的境地，因为各州不情愿也没有能力投入这笔资金。假定州政府交通部门愿意承担V2X系统的费用，另一个巨大的障碍是缺乏安装、维护、运营V2X技术设备的人才。

V2X技术的市场化是个公认的难题。普林斯顿大学的机器人研究教授阿兰·肯豪森（Alain Kornhauser）是资深的自动驾驶研究人员，他不赞同美国交通运输部安装V2X。肯豪森写了一封关于自动驾驶最新进展的邮件，展现了他对这个问题清晰的洞察和一针见血。在邮件中，肯豪森以标志性的评论（“得了吧”“这些人根本不带脑子”）戳穿了那些愚蠢的新闻故事。肯豪森也对“仅在新车安装V2X，其他车辆并不连通”这种做法提出严厉的批评。他指出：

如果交通运输部固持己见，要求所有新车安装V2X，问题并不会有太大改善。V2X需依赖大多数车才能起作用。也就是说，

V2X的传感器只能在其他车辆也匹配了V2X的发射器情况下，才能有效工作。只有当道路上的车辆超过一半安装了合适的软件，“在新车上安装V2X系统”的规定才能真正受益。

想要提高基础设施的数据传输效率，市场化是关键。正如肯豪森指出的，假如道路上仅有10%的车辆安装了V2X，那么只有1%对（ $10\% \times 10\%$ ）车辆能够实现信息交互。也就是说，在全美2.5亿辆行驶汽车中，尽管最终达到有10%安装了V2X，也仅有1%的交通事故出现在两辆数据互通的汽车之间。假设这些汽车的人类驾驶员精力集中，对预警提示反映得当，那么这样的V2X安装率所能降低的事故概率仅不足1%，可算不上是多大的进步。

接下来，挑战接踵而至。目前，全美尚未建立起统一的V2X技术数据传输标准。缺乏公认的国家级数据标准可能导致V2X汽车仅限于当地城市、所在州道路传感器之间交互信息。大部分的驾驶路线都会经过不同的城市和州界，基于这一客观事实，地区之间不同的V2X数据标准极大限制了这种全国性基础设施的实用性。

事实上，这还没有算上复杂多变的人为因素。V2X的研究具有典型的“脱离时代”的特征，人类司机自始至终都需要在其中扮演关键的角色。区别于全自动化汽车——人工智能系统的驾驶技术有望超越人类，但V2X技术只能单纯发出警示而已。如果人类司机对警示反应不及时，V2X的安全效益便等于零。更糟的是某些时候，“哗哗”的声音容易分散司机的注意力，搞不清楚状况；最坏的结果是，司机因为提示而走神，反倒忽略了眼前最重要的事情：保持安全驾驶。

最后，像所有的点对点（peer to peer，P2P）无线网络一样，V2X系统对防御黑客入侵、系统干扰和诈骗等风险能力极其微弱。软件行业再三汲取的深刻教训是，任何分散形通信机制都无力抵御黑客入侵和破坏的风险。任何制造系统都存在漏洞和不到位之处。无论有多少专家学者、工程师团队参与设计，研制出的交互系统多么“万无一失”，总会有些“防不胜防”的人。V2X基础设施一天内需要顺利完成上百万辆个人轿车之间的上千次交互，而黑客入侵系统只需要一次得手便能大肆破坏。同时，常见的通信干扰器便能轻松干扰V2X信号。技术娴熟的黑客只要模拟出可靠的信号源，便能制造混乱。假设一辆车发出虚假信号“我以每小时100英里驶向十字路口”，附近的V2X汽车系统便会“哗”声一片，导致车内司机陷入恐慌。

基于以上缺陷，为什么专家们还会继续追求发展V2X技术呢？V2X，这种车联网技术极具讽刺意味的一点是，只有当道路上行驶的每一辆车子都全自动化以后，车联网才有意义。在资源无限、车辆全自动化的理想环境中，车-道路基础设施数据交互才能带来若干好处，比如实施交通流量优化、紧急车道优先与道路状况警示。

在过去几年内，设备连接的数量不断激增，使得物联网的概念在技术界炙手可热。高德纳咨询公司（Gartner）发布的一份报告中预测，在2016年，智慧城市中联网设备的数量高达16亿，同比增长39%。到了2018年，联网设备的使用数量上升至33亿。大部分的联网设备用于安防（比如摄像监控），在商业建筑或公共空间（大型商场、办公园区和机场）内调控气候。在家庭中，联网设备用于小型的娱乐设备、安防、温度、通风控制等等。

此时，让我们想象一下，当道路上大部分汽车达到100%无人驾驶和自

动化，无线传输器被普及到每个路口和高速公路入口。我们再把学校区域和建设施工场所加入网络中，然后假设所有的车辆都安装了完美的防误、防黑客无线发射器和接收器，于是发射器、交叉路口和学校区域三者之间流畅地共享交通图、道路突发状况等数据。

以上的情形包含许多“假设”的理想因素，但为了达到预期目标，还得继续。想象每辆自动驾驶汽车里的交通控制软件都能成功接收周边基础设施发出的信息，然后启动车速和驾驶线路的智能化调节。于是，一切发动机的空转都将被拒绝。

自动驾驶汽车从道路的物联网上获取信息，据此精准规划驾驶路线。从此我们与交通堵塞告别。由此也能带来效益巨大的安全性提升，因为三分之一的夺命事故都发生在各个交叉路口。

互联智能汽车之间能在彼此间交互数据，形成中心明确、高效整合的信息体系。当年宝马汽车收购HERE数字地图，预测收购的益处如下：

集群智能带来巨大的社会效益：以冰封路面为例，通过启动ABS防抱死制动系统，同时根据外界温度，顺利检测突发状况并实时预警。将来，我们能更精确地辨识交通拥堵，极大降低事故风险。当系统感应到绿灯亮起时，为广大车辆提供信息导航，帮助其选择最佳油门，节油环保。

未来，互联智能汽车数据共享的实现不是梦。但到那时候V2X汽车不太可能仍旧使用美国交通运输部目前政策中的短波无线电技术。等到全自动

化汽车普及后，互联成本将会下降。因此，没有必要投资在天价的短波无线电设备上，车辆在交通管理服务器的帮助下彼此交通，而服务器可以直接使用现成的手机通信设施。

非智能高速公路的价值

或许是由于近几十年间，联邦政府和私营部门的资金不断投入到智能高速公路系统领域，很多人觉得无人驾驶汽车上路，需要对高速公路基础设施进行大规模的投资。直到2014年时，沃尔沃公司还提议把磁铁嵌入公路，以便无人驾驶汽车可以在雾中辨识道路。至此，改进高速公路基础设施的建议势在必行。就连那些资深的计算机程序员也十分渴望，设想着若是能实现条形码在公路上的应用，或是把射频识别系统（RFID）安装到每个交叉路口，那么无人驾驶汽车的编程将会简单得多。

但鉴于诸多原因，投资智能高速公路被认为是不明智的。其中一条非常现实：没有用于非核心的高速公路基础设施项目的现成专项资金；而申请拨款会使该项技术发展陷入被动的局面。在公共道路破损坍塌、高速公路信托基金几近破产的背景下，在美国为高速公路基础设施筹资一向被视为棘手的政治难题。即便筹资这条路行得通，把越多的资金用在高速公路基础设施建设上，可投入在飞速发展的无人驾驶技术（机器人可比人类更擅长驾驶）上的资金就越少。

另一个反对投资建设智能高速公司的理由如下：既然汽车能做到安全

驾驶，便不再需要交通标示、红绿灯和防护栏。对于无人驾驶车辆行驶而言，最重要的道路基础设施也不过是技术含量不高的车道标志，能为汽车的系统软件提供关键的视觉信息即可。可车道标志的视觉效果却因各个州政府而异，差距不小。无人驾驶车辆的支持者应当集中政治资本，促进车道标志制作的政府标准早日出台，由此不难预测，一场混战将在各个自治市之间、郡与郡之间、州与州之间展开。

第三，由于硬件技术很快就要面临被时代淘汰的命运，智能基础设施的投资风险日趋增大。通常来说，软件的更新换代比配套硬件速度快，周期短。因此，宝贵的研发资金的最佳投资领域要数无人驾驶汽车的自动操作系统，因其性能会不断提高，甚至实现“裂变式”增长。

“非智能基础设施”才是我们的最佳之选。美国奇点大学的自动驾驶研究专家布拉德·邓普顿（Brad Templeton）认为，“互联网”是一个很好的例子，凸显了非智能基础设施的价值所在。互联网基础设施本身固有的智能特性甚少，不像人们想象的那样，拥有自适应性和自动升级的功能。路由器和网线能传输数据包，根本不了解数据包里面的内容——管他们是电邮、视频还是网页呢！正是由于互联网的构造趋向尽可能地非智能化，才推动了技术畅通无阻的发展。

假设在一个虚拟的世界中，网络设备全智能化，能够充分了解数据包的内容，进行路径传输和策略处理。在这样的情况下，软件应用每变化一次，网络设备也要随之升级。

同理，无人驾驶汽车也一样。交通设施越简单，政府机构的参与便会越多——具有自由，灵活和适应性强的特点。邓普顿投资公司把智能汽车

与铁路运输做比较。铁路的轨道要比汽车走的公路要更特别和单一。公路的用途广泛，能够承载多种交通工具，从单车、行人、牲口到小轿车、货车都能在路上走。公路显然能承载更复杂的交通形式。司机可以自由变换车道，而铁路的造价昂贵不说，且一次只能通行一辆火车。

修订交通运输政策

十几年前那场轰动一时的“Demo 97”活动后，美国交通运输部与其下属机构高速公路安全协会和联邦公路管理局喜欢以消极被动的态度发展无人驾驶汽车，置身事外看看是否有行业或哪个州愿意承接这个项目。部分联邦官员认同某些保守的做法，例如对信息与通信技术的发展不予评价。然而，无人驾驶汽车的技术发展日新月异，使得当局再也没有借口仅局限在V2X研究上。

同时，也有一些令人欣喜的迹象表明情况正在改善。2015年12月，国会通过一项人们翘首以盼的法案——《美国道路交通修订法案》（*Fixing America's Surface Transportation Act*），大力推动了联邦高速公路基础设施和传输系统的建设。根据该法案规定，在5年内将投入约3000亿美元用于道路和传输系统建设，并从国库拨款用于发展先进的交通管理技术；同时要求财政部向国会递交关于交通自动化政策的详细报告。

2016年，美国交通运输部部长安东尼·福克斯（Anthony Foxx）宣称未来十年内，投入将近40亿美元在某些试点项目上，以加快自动化车辆的发

展和应用（截至撰写本段落，我们还不清楚所谓试点项目的发展重心是什么，希望不仅是V2X技术）。这次，交通运输部采取了较为积极主动的态度，为自动驾驶汽车安全操作规范提供指导意见。交通运输部将与各州联手，出台自动驾驶汽车测试与应用的规范性政策的草案。

如果我们能跻身进入联邦政府内部，我们该做些什么来加速无人驾驶技术的发展呢？首先，设立专门管理全自动化汽车的机构，将其命名为联邦自动汽车管理局（AVA），类似于美国联邦航空管理局（FAA）负责管理航空秩序，联邦自动汽车管理局的职责是为无人驾驶汽车制定积极可行的政策，让无人驾驶在全美50个州成为现实。

航管局为飞机的数量和检测制定了严密的衡量标准，无人驾驶汽车同样也需要一系列的严格监管。联邦交通局的官员们需要负起领导者的角色，为某些关键问题准确定义，例如无人驾驶汽车在什么情况下才能被称为“安全有报障”？举个具体例子，无人驾驶汽车平均行驶50万英里的距离允许发生一起碰撞事故。达到这样的标准（这可比出租车司机的驾驶门槛还要高）是否可以在公共道路上行驶？安全标准一旦确定后，就要马上去落实。

安全问题讨论的另一方面是许多人难以接受无人驾驶也有风险、不确定因素，甚至发生事故。问题就在于人类司机更不可靠，且有大量记录可以佐证。监管人员不能坐等无人驾驶汽车自动变得完美，因为那天也许永远不会到来。联邦的监控者应当为安全性划定新的实用的范围：正如之前讨论的，无人驾驶汽车的安全标准应当明确，为人类司机的两倍——平均每40万英里有一起碰撞事故。

另一方面，自动汽车管理局虚拟机构应当致力于推广无人驾驶汽车，让群众和公司产生兴趣、予以关注，才能鼓励更多的人投身解决现有的技术问题。汇集民智，共襄盛举，这不仅仅是谷歌、苹果和各大汽车公司职员的事情。为此，自动汽车管理局应当拨款举行一年两期的知识竞赛，组织来自各行各业和高校的团队参赛，解决所面临的技术难题。如果说每年一度的无人驾驶知识竞赛听起来有些熟悉，那是因为这种形式套用了之前的DARPA 2004年、2005年、2007年的大挑战，其间催生了许多良策，帮助无人驾驶技术蓬勃发展至今。

各州和各城市都需要国家予以指导。自动汽车管理局应当下令各州提出一套无人驾驶驾照标准方案，让该项技术能通过起码的安全门槛。2015年只有四个州颁发了无人驾驶汽车驾照：加利福尼亚、内华达、密歇根和佛罗里达，这些州鼓励把州境内足够长度的、闲置的高速公路辅道用作测试、确认全自动化汽车的达标情况。

以上目标无一能轻易完成，制定良好的技术政策更是众所皆知的艰难。技术政策需要顾及汽车和安全性能，更要在加油鼓劲和鲁莽冒进之间找到平衡点，这实在是个挑战。最近，加州通过了无人驾驶汽车的修订案，为尚处于探索阶段的全国层面监管照亮了茫茫前路。

2015年12月16日，加州机动车管理部门（DMV）根据公众反馈，发布了首份无人驾驶车辆监管条例草案。条例最具争议之处在于要求无人驾驶车辆必须同时配备方向盘、刹车系统和人类司机在车内，方能上路。要求的不合理之处在于没有充分考虑其中的风险——让人类持续地监管高安全性能的无人驾驶车辆。当人类相信计算机系统能把一切都处理好的情况

下，他们绝不会用心驾驶。该要求的另一个弊端——“人类司机必须在场”的奇怪要求让无人驾驶陷入“必须有人”的怪圈。这种方法与谷歌公司研发全自动化汽车的想法南辕北辙。依据加州监管条例的规定，谷歌公司对无人驾驶汽车的设想——黑色按键开车，红色按键停车——显然是违法的。

如此考虑不周的条例一旦通过，就难以收回。布拉德·邓普顿总结了该条例可能产生的不良影响：

加州提出的禁止谷歌风格的无人驾驶汽车，将来会通过法律形式来落实。不幸的是，某种事物一旦被禁止了，想要解禁就难上加难。因为多一事不如少一事，没有人想给自己揽上监管的活儿；政客更不情愿冒着被人责骂的风险解禁某种事物。无人驾驶汽车也会带来伤害，只是比目前人类驾驶所带来的伤害要少。

不过，该项新监管条例中也开创了好的先例。加州要求无人驾驶汽车的生产商纳入第三方进行车辆安全测试——截至撰写本书之时，测试体系尚未成型，使得该条例“实际操作困难”——新监管条例的合理之处不仅仅在于推动一项新行业（无人驾驶汽车测试）的发展，同时要保障消费者免受山寨汽车软件和虚假承诺的危害。我们应当进一步建立健全监管条例，清晰定义政府对个人自动汽车的监管办法。我们会定时检查车子，避免刹车失灵和尾气超标。与此类似，无人驾驶汽车也需要周期性的检查来保障软件、硬件的顺利运作。

自动汽车管理局应当指导各州定义和规范另一个安全相关的重要问题——划分责任。无人驾驶汽车发生事故时，过错由哪一方承担？由于无人驾驶汽车发生车祸相对罕见，这个问题很可能就此不了了之，但事实上我们需要客观审视它。美国保险法的法律效力适用于全国，如果联邦政府能全面针对无人驾驶汽车设立清晰的法律标准，覆盖软硬件传感器、汽车车身等各个方面，那么保险公司可以明晰风险，生产商也能避免遭遇恶意的诉讼。

在对各州和各大城市进行监管时，联邦当局还有一项需要认真考虑的事项——隐私保护。无人驾驶汽车的记录涉及乘客大量的出行数据，汇集了众多公路可视性的数据。截至目前，我们已经看到一些保护隐私方面的初步措施。2015年纽约州议员孟昭文（Grace Meng）提出了保护无人驾驶汽车消费者隐私的法案，我们朝正确的方向迈进了一步。

我们希望美国交通运输局尽早加强对无人驾驶汽车的全面监管。制定万全之策的重点在于，一方面加快出台有利政策，为无人驾驶汽车发展铺平道路，更好地保障公共安全；另一方面，帮助各大城市、各州实现无人驾驶汽车的普及。有前瞻性政策的指引，各地区才能在将来收获经济红利的同时，成为无人驾驶技术发展重心。

如果无人驾驶汽车的发展在“Demo 97”活动之后停滞不前，那么本书就不会存在了。在前面简要的历史回顾中提到，自20世纪以来，历代发明家致力于研究可推广的无人驾驶汽车，他们不断尝试却屡屡失败。原因在于缺乏核心的支持技术、系统深度学习能力、数据整合、高速计算机以及先进的传感器。令人欣喜的是这一切现在都依旧在高速发展。

21世纪初，现代无人驾驶汽车诞生在一片荒芜的沙漠里，远离电子高速公路和煤渣砖块铺成的小道。它不受高速公路设施的束缚，并取得重大突破——把智能设备安装在车本身的轮式移动机器人。这种新型的无人驾驶汽车就是机器人。

第七章

机器人的崛起

Rise of the Robots

现在看起来，第一批现代无人驾驶汽车非常简陋，几乎是一群“视障”机器人。一旦视觉系统不能识别某样物品，就会畏缩不前。20世纪八九十年代，德国自动汽车先驱恩斯特·迪克曼斯（Ernst Dickmanns）制作出几辆无人驾驶原型车。这些汽车使用智能汽车和传感器来辨别方向。同时，意大利的阿尔贝托·布罗基教授（Prof. Alberto Broggi）也创造出一辆车子，能够根据视觉软件的指引，沿着公路标示行驶。尽管用今天的标准来看，早期的无人驾驶汽车非常原始，可是与五六十年代由无线电引导的别克车相比，它们有着巨大的优势：其智能化在车子本身，并非在公路上。

现代无人驾驶汽车构想的实现归功于两大事件：第一，微处理器体积变小，而承载的功能却更强大了；第二，2001年美国国会通过一项法案：2015年前，美国有三分之一的军用车辆必须使用全自动化汽车。就像手机、GPS导航系统和互联网一样，无人驾驶汽车原本服务于军队，后来进入了日常消费者的生活。

根据该法案，国会命令DARPA负责推动相关技术的发展，授权该局对无人驾驶汽车的研发人员进行现金奖励。有了奖金的支持，DARPA当局制定了相应计划，赞助了一项公路比赛。研发人员让自己的机器人汽车与来自其他高校、汽车企业的车辆展开比赛，竞逐现金大奖。

（读者们回想一下，在这以后，DARPA还使用了类似的方法来推动灾难般的CHIMP机器人的发展。）

2001-2007年六年间，DARPA分别赞助了三场公路赛事——2004年、2005年与2007年的DARPA挑战赛。2004年首届DARPA挑战赛，冠军团队的现金奖励高达100万美元，非常诱人。哪家的自动汽车能成功通过位于美国西南部的不毛之地莫哈维沙漠的150英里长赛道，就获胜。沙漠是第一届比赛地点的首选，由于当时的无人驾驶技术非常不完善，机器人专家必须选择远离繁华的地区进行测试，远离那些购物中心和熙熙攘攘的街道，避免潜在的风险。

事实证明，选择在沙漠开赛是正确的。由于车载软硬件系统过于粗糙，15支参赛团队都没能顺利完成眼前的任务。硬件传感器和GPS系统速度慢不说，而且极其不稳定。机器的软件表现就更糟了，致使车子在路堤和岩石上停滞不前。参赛车辆由于各种机器问题纷纷退出比赛，比赛进行了几个小时以后，15支参赛队伍中竟然无一能行走超过8英里的路程。最终，这一届比赛的100万美元的现金大奖没能送出。

面对如此惨淡的比赛结果，几乎每一位参赛者都会冒出放弃的念头。2004年挑战赛结束后，CNN对DARPA挑战赛的项目副经理汤姆·斯特拉（Tom Strat）进行了采访。汤姆本人却依然很乐观：“尽管没有一位选手能

完成超过5%的赛程，但这更将激发工程师们的斗志。”

DARPA并没有知难而退，继续为下年赛事的举办筹集奖金。2004年的比赛汇集了各式各样令人眼花缭乱的自动化汽车，从小型承重2吨的卡车到车轮巨型的沙漠越野车，无奇不有。这次以后，在2005年的比赛中，主办方提高了参赛选手的门槛，对比赛场地严格把关，专门举办了一场全国资格赛，而冠军车队的奖金增加到了200万美元。

根据各参赛车队的资格赛成绩，DARPA从中选择了前23名参赛者进入2005年的DARPA挑战赛。比赛场地仍然是荒芜的沙漠，规则与2004年大致相同：参赛车辆在不借助公路设施和外力帮助的条件下，自动驾驶通过132英里长的越野赛道。

很遗憾，只有在半个世纪前任职通用汽车公司和美国无线电公司的工程师（比赛的评判者）才有机会真正亲眼见证这空旷沙漠里的奇迹——2005年的DARPA挑战赛是移动机器人发展史上的临界点。史上第一次，五辆无人驾驶汽车使用人工识别系统，成功通过了路况恶劣的沙漠赛道。

2005年挑战赛的冠军是斯坦福大学车队，他们在七小时内顺利完成了比赛。卡内基·梅隆大学的两支车队紧跟其后，分别获得亚军、季军。第四名由来自格雷保险公司（Gray Insurance Company）的选手摘得，第五名则花落美军重型战术卡车的唯一供应商奥什科什卡车公司（Oshkosh Truck Corporation）。

斯坦福车队赢得冠军的过程，比挑战赛结果更让人兴奋。有些车队使用拓扑地图（Topographic maps）和航空影像生成车行线路，而冠军车队，

一辆名为“史丹利”的改装版大众途锐则另辟蹊径：史丹利的中级控制系统使用机器学习软件，成功地做到了识别与响应。

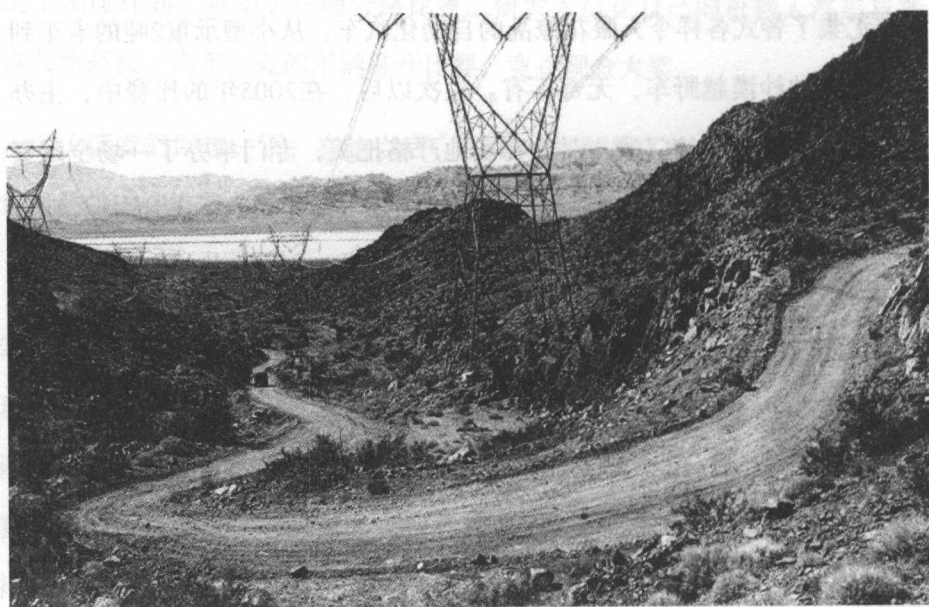


图7.1 2005年DARPA挑战赛中，位于内华达州路西·格雷山（Lucy Gray Mountains）的Beer Bottle Pass赛道，此处距离终点约有7英里，颠簸崎岖的高难度路段超过20处

来源：维基百科，DARPA挑战赛（2005年）

机器学习和无人驾驶

斯坦福大学教授塞巴斯蒂安·特伦（Sebastian Thrun）带领那支研发出“史丹利”的冠军团队，对赛前准备有着不同于对手的思考。首先，他意识到是汽车软件，而不是硬件决定比赛胜负。其次，是对中级控制系统软件的研发。考虑到汽车引导软件中“识别”和“响应”两大功能的重要地

位，他们摒弃了当时盛行的中规中矩的基于规则的人工智能软件。在项目初期，他们也曾尝试去写一套符合逻辑的指令，指导汽车在比赛过程中处理可能遇到的障碍，事实证明这种做法行不通。最终，特伦与其团队选择了机器学习的方法，他解释道：

许多参赛者非常关注硬件，于是许多车队研发出独家的机器人汽车来参赛。我们则认为，比赛重点不在于机器人的力量强弱或汽车底盘的设计优劣。如果是人类驾驶的话，任何车都能够顺利通过沙漠赛道，这段越野赛道并不复杂，只是一段普通的沙漠车道。于是，我们决定仅研究人工智能，把一台计算机安装在车内，为车子安装“千里眼”和“顺风耳”，打造出一辆智能化汽车。

在汽车智能化的研发中，我们发现简单的规则不足以指导软件完成驾驶，需要成百上千条代码指令，才足以应付多种的意外情况。某天的测试中，公路旁有一群鸟。当汽车靠近时，鸟群飞起。这时候我们才发现机器人的“眼睛”里无法识别鸟类和岩石！于是我们不得不提高汽车智能化程度，把鸟和石头区别开来。

后来，我们依靠所谓的机器学习或大数据来驱动汽车：与其写出所有的程序，不如用教人类驾驶的方法来教机器人。我们来到沙漠，我开车，机器人“观看”并“模仿”我所有的动作。后来，我们直接让机器人来驾驶，当它犯错的时候，我们返回到数据中解释错误原因，给机器人机会改正错误。

要搞清楚为什么特伦的机器学习在当时被称为一项重大创新，一起

来回顾前面几章提到的两种主流人工智能软件类型：自上而下的基于指令的符号型人工智能和自下而上的数据驱动型人工智能。正如我们之前所讨论的，自上而下的符号型人工智能要求程序员必须先搭建出整体的理论模型，然后写出一系列应对各种状况的指令，与模型产生交流。不同的是，自下而上的数据型人工智能，例如机器学习，采用大量的数据算法并使用相应的技术进行数据处理，最终使得汽车软件无需人类监管，也能自主学习并“识别”固定的模式。

在第四章的一项实验中，假设现在要利用中层控制软件来指导无人驾驶汽车通过一个繁忙的十字路口。在这项实验中，我们发现无法通过写代码的方式指导汽车处理每一种状况，或定义每一处在现实世界中可能遇到的障碍。仅仅依靠计算机软件发出死板的“如果-则”指令，无法灵活调控汽车的视觉和响应功能。用不了多久，临界状况或指令以外的小插曲就足以令它寸步难行。

现在设想进行一项类似的实验，不过这次实验的目的是利用中层控制软件来指导无人驾驶汽车穿越沙漠，那么这种中层控制软件必须具备强大的识别功能，准确分辨车辆前方路段能否安全通过。一个可能的办法是利用航空影像和GPS定位系统为参赛车辆生成路线，穿越沙漠（2005年DARPA挑战赛中部分参赛者使用了这种方法）。经过考虑，我们认为该方法治标不治本。当车辆在行驶途中遇到系统预设好的指令以外的情况，如凹凸不平的路面、坑洞、岩石和沟渠等，软件将不知所措。

于是，我们得出一项重要结论，为了确保无人驾驶汽车的安全、顺利驾驶，需要构建中层控制软件实时读取从传感器输出的路况数据。在此

基础上，仍然有许多程序员倚赖死板的指令方法来定义“驾驶功能”。在2005年DARPA挑战赛中，某支参赛队伍花费数月时间，编写整套逻辑指令来处理传感器输出的数据流。他们希望当车辆前方的路况数据增长时，系统控制软件能够指示车辆转向，绕开障碍物。经过数月的艰苦努力，他们创建了数量巨大、详细的数据库。遗憾的是，这支车队赢取百万奖金的旅程停止在比赛途中。他们的车辆由于制动不及时，在隧道入口发生碰撞。对此，该车队表示他们在预先编写中，缺乏应对隧道的特殊指令，团队在赛前准备时没有预料到这种状况。在缺乏清晰指令的情况下，系统只能根据数据胡乱猜测：隧道顶部很高，且赫然出现在车道中央，于是系统软件将其判断为一座巨大的高墙。在这样的误解下，系统作出自认为对的处理：踩下刹车，拒绝前进直等到工作人员到来，将它带回安全地段。

假如我们是参与“史丹利”项目的研发人员，就会明白特伦团队的选择：自动驾驶不是两三条指令可以做到的事，需要成百上千条指令方可成事。指令数量之大，绝对不是凭几个程序员的力量可以完整写出来的。哪怕在一条空旷的赛道上，也可能发生无数人们始料未及的新状况。以指令为基础的符号型人工智能，只能在沙漠路面结冰的时候起作用。配备测量团队和研发团队投入无限时间精力，创建出覆盖整个路面的数字地图，详细到路面每一块小石子和车库里一点点的不整齐之处。哪怕人们创造出了这幅完美的数字地图，该模型的代码基础仍然不能避免失败，因为客观世界会出现各种混乱的突发状况，千变万化。

以指令为基础的软件价值体现在无人驾驶汽车的高级控制功能应用上，例如路线规划、检查油箱等简单的操作。然而，由于这种类型的人工智能在非结构性环境中很容易崩溃，所以也给许多机器人专家留下“脆弱不稳”的

印象。在2004年的DARPA挑战赛恰好反映了这种情况，参赛者使用的软件无法持续正常工作，结果全部失败，最好的驾驶成绩是8英里路程。

第三届，也是最后一届DARPA城市挑战赛在2007年举行。在这届挑战赛中，自动汽车收获了好几项突破性的成果。为了让各参赛队伍能够有所突破，主办方将比赛地点选在洛杉矶东北部75英里的一座废弃空军基地，赛道长60英里。该赛道状况无法事先预测，充满未知数，模拟的是在混乱的战争地区或车水马龙的高速公路上行驶的场景。想要赢取冠军的200万美元奖金，各参赛车队需要在没有清晰指令的条件下，通过中层控制软件在陌生的交通环境中或高速行驶的车流中安全驾驶。在当时，要完成这项任务，面临的挑战就好比人类在伸手不见五指的暴风雪中攀登珠穆朗玛峰，并且没有地图可以参阅。

2007年DARPA挑战赛规则简单明了：谢绝人类司机在旁，车辆要在城市交通的环境下自主完成一系列简单的任务。任务包括，左转行驶至十字路口，通过环状交叉路口，停车，在两车道的公路上持续平稳驾驶，不发生交通事故。为了确保参赛车辆是真正的全自动化（并非事先根据场地设写好相应程序），各参赛队伍在赛前一小时才拿到各自的参赛路线文档，这相当于人类的司机不能提前熟悉赛道。

2007年度挑战赛开赛那天，来自各大高校和汽车公司的11辆参赛车辆在发车线那头跃跃欲试，百余位各车队的工作人员也在发车线附近摩拳擦掌。当头戴棒球帽的发令员一声令下，挥下手中的小绿旗，比赛开始！无人驾驶汽车一辆接一辆从车门处平稳开出。只见计算机被安装在车辆的车尾和后座部分，方向盘来回旋转，仿佛有一只无形的手在掌舵。离赛道不

远设有一处巨大的帐篷，上千名无人驾驶车爱好者以及观众在一块巨大的屏幕上关注着车辆的一举一动。

该年度的比赛过程中出现多种状况。裁判身穿橘红色反光安全马甲，手持计时器，看着各大车辆在酷热的太阳下穿行于南加州沙漠，完成各自任务。车辆则看起来像视力衰退的八旬老人组成的队伍，一辆接着一辆缓慢前行，绕过主办方设置的混凝土路障。

在这场世界级精英机器人的角逐中，2007年的无人驾驶汽车技术仍然让人不放心。为了维护比赛秩序、确保裁判和参赛者安全，每辆无人驾驶汽车都会配备一位“保姆”——一辆由人类驾驶的加强版福特金牛座（Ford Taurus）。万一无人驾驶汽车内的人工智能系统失灵，“保姆”上的驾驶员会马上远程启动无人驾驶汽车内的紧急停止装置，避免对附近的人车造成伤害。

这条沙漠赛道，曾经是二战和朝鲜战争时的美国空军的训练场，现在看起来就像一幕生动的慢速播放的、讲述车辆间发生小摩擦的喜剧电影。麻省理工学院的参赛车辆向康奈尔大学的车辆渐渐靠近，结果发生了我们在前面第四章提到的交通事故。另一辆参赛汽车也很快结束了自己的比赛，它任性地偏离赛道并撞上了附近的建筑物，让整个比赛现场变成了“车辆与建筑物碰撞”的事故聚集地。另外两辆汽车就像初次登台表演的青少年，畏首畏尾，在原地停滞不前，仿佛在思考应该朝前方的十字路口还是朝环形交通路口前进。

尽管出现了不少突发状况，比赛的最终结果还不错。11支参赛车队中有6支完成了规定任务并跑完全程。冠军由卡内基·梅隆大学和通用汽车联

手研发的“老板（Boss）”汽车获得，全程用时4小时10分钟，平均速度每小时14英里。斯坦福大学的“青年（Junior）”紧随其后，获得亚军。弗吉尼亚理工大学的车辆“奥丁（Odin）”获得季军。

然而，那一天真正的胜利者是关于机器人的学科，2007年度挑战赛的赛果证实无人驾驶汽车完全可以在四车道的公路上自动驾驶，并能检测识别路上的其他车辆。经过将近一个世纪的“解放双手，解放双脚”研发工作，自动化汽车终于可以正常上路了。后来，困扰了无人驾驶汽车发展数十年之久的达·芬奇难题终于要得到解决。

西洋跳棋

有人指出2007年是现代全自动汽车的诞生元年。事实上，现代无人驾驶汽车的出现分为不同阶段。早在2004年、2005年举办的前两届比赛中，无人驾驶技术的表现就开始渐入佳境了。到了2007年的第三届比赛，参赛者不仅受益于之前的大赛经验，帮助提升比赛成绩的还有硬件技术的迅猛发展、以机器学习为代表的人工智能软件的重大突破。

Stack Overflow是一个著名的IT类网站，来自全球各地数百万的程序员在网站上交流探讨技术问题，并根据解答质量的优劣对回答进行排名。在Stack Overflow上，我们找到了对机器学习定义的解释。“机器学习是什么？”该问题排名最前的解释如下：

从根本上来说，机器学习是一种计算机的教学方法，用于指导计算机培养或提升识别能力和基于数据产生的行为。这里的“数据（data）”是指什么呢？这要视具体情况而定。可能是机器人学习行走时，传感器的数据读取，也可能是某个项目数据的正确输出。这种对输入数据前所未有的反应能力是许多机器学习算法的核心原则之一。假设现在要教会一台计算机在高速公路上安全驾驶，倘若使用数据库方法，你必须教会计算机处理上百万种情况。而高效的机器学习算法能够在不同情形中找到相似性，然后做出相似的反应。

不同情形下的相似性可以是任何东西，甚至我们能想到的日常生活中的例子，都值得机器学习和模仿。例如，当计算机驾驶系统学会了“当前面的车辆减速的时候，车辆本身也要跟着减速”。对于人类而言，把汽车替换成摩托车，我们也会采取同样的处理方式，因为我们知道摩托车也是机动车的一种。但是机器学习算法直接处理这种情况的困难程度可能会令人惊讶！数据库里需要分别储存路面前方是汽车或摩托车两种情形。不过另一方面，机器学习算法能够根据汽车的例子自主“学习”，生成摩托车的处理方式。机器学习的另一个特点是模式识别，教会某种程序对多种模式作出反应或辨别。

机器学习技术听起来自然而然——软件自主学习模式以识别和解决某些问题——事实上发生的是：一个从海量数据中统计并寻找规律的分析算法。该算法使用已有数据模式来搭建数学模型，数学模型会把各种各样的

输出结果根据可能性高低排行，最终得出预判或结论。接着，算法会把这些结果应用到新的、尚未正式的数据上来测试，核实之前的预判是否准确（决策是否合适用）。如果出现错误，就返回更新模型。这样一来，机器学习项目就能提供从“学习”到“体验”整个过程的数据，并且这一切都在项目管理人员的监管下进行。他们的工作内容包括选择相应的算法、提供数据和反馈正误。

棋牌游戏一直以来都是人工智能研究人员，展示新逻辑技术时最喜欢的方式，机器学习也不例外。在20世纪50年代机器学习技术的发展刚刚开始起步，计算机有限的能力极大地限制了可供选择的桌游应用范围。由于那个年代的计算机还不能处理国际象棋所用的计算，研究人员于是用西洋跳棋（checkers）来代替。

1949年，IBM公司的新员工阿瑟·塞缪尔（Arthur Samuel）想要证实计算机能执行复杂的智能任务。塞缪尔入职那年，IBM还只是一所主要以生产老式计算机著称的公司。塞缪尔对于如何提高公司的知名度很有自己的想法。他认为，如果能研发出某种只适用于计算机，而不适用于老式加法机的应用，那么他就能展现IBM第一台商业计算机IBM 701的高超分析能力。

再没有比西洋跳棋更适合证明计算机的认知能力的方式了。塞缪尔的目标是提升计算机的西洋跳棋水平，达到世界大师的级别！如果选择用当时主流的人工智能模式来解决问题，他需要写出大量的计算机程序指令用来预测、指导计算机处理可能出现的棋盘布局。

根据经验推测，使用这种规则库方法将会是一项艰巨的任务，因为每一种可能出现的棋盘布局，都需要预先设置好该情形所对应的规则。应用

情形如下：一项规则指示“先去吃掉对方的棋子”，另一项规则指示“先移动靠近对手后方的棋子”。

塞缪尔很快察觉这类人工智能的缺陷，下好一盘棋需要使用大量的数据，还有一串长长的难以管理的指令名单。更大的问题是，就算有人给每一处棋盘位置都列出对应的规则，计算机再怎么跟着这些死板的规则来下棋，最终也只能成为一个中规中矩的棋手。这就好比一个初出茅庐的新手，仅仅接受刻板僵化的战术指导，培养不了“棋感”，计算机也是一样，它对棋局缺乏全局性的洞察力和一个伟大棋手身上的特质。

当时许多人工智能专家的研发都取得了很大进展，他们通过编译一系列复杂的指令，刻苦钻研意外案例和“伪随机性（Psuedorandom）”的下棋点，这样计算机看起来似乎运用策略在下棋。塞缪尔却另辟蹊径，他选择了机器学习的方式。因为只有这种方式，计算机才真正以自己的下棋经验为基础学习棋艺，而不是跟在一堆指令后面亦步亦趋。

人类棋手的成功之路不是从计算各种可能性的输出开始的，而是通过观察和记忆典型的棋局，记住那些精髓。人类棋手会牢牢记住关系成败的那一步棋。塞缪尔决定让计算机学习资深棋手，边下棋边学习如何识别典型的棋局模式，尤其是那些经典制胜棋局。

塞缪尔让计算机从随意乱走开始，并拷贝了一份同样的软件作为其对手。有时候原始版本的计算机程序会赢，有时候会输。每一局结束后，计算机把所有的数据存储在数据库里，无论是胜局还是败局的下法。这个过程就是计算机积累实战经验的过程。

迎战下一个棋局前，计算机已经有了相对丰富并持续增长的实战经验，这样软件系统在每走一步之前，会在数据库里找到对应的棋局。通过这种途径，它能找出之前是否遇到相似的棋局，如果曾经遇到它能顺利找到通往胜利的下一步棋。如果软件从未见过这种棋局，软件会随机地按照下棋规则移动一步棋，并把那步棋的结果才储存下来。

刚开始塞缪尔的计算机随意地下棋，就像小孩子懵懂地开始下他人生的第一盘棋。下了上千局棋以后，数据库存储了丰富的各式棋着。经过这些棋局之后，计算机才形成了自己的方法——某些观察家口中的“策略”。多数情况下，走一步棋并不能马上定胜负，成败是由持续的走步策略逐渐决定的。因此，数据库不仅仅记录每一步棋的“输/赢”输出，它记录的是每走一步通往成功的可能性大小。换句话说，数据库本质上是一个巨大数据模型。软件学到这个程度以后，它会在自己和自己对弈中投入大量的时间，积累了人类棋手一辈子都不可能获得的、数不胜数的实战经验。

随着数据库的扩张，缪赛尔必须研发出更为高效的配套数据检索方法，比如至今仍在大型数据库里应用的散列表法（hash table）。缪赛尔的另一项创新是利用数据库来推算出对手的下一步棋会怎么走，这就是今天的极小化极大算法（minimax）。

最终，缪赛尔成功了。他的计算机下棋程序走出实验室，对全世界造成了非同寻常的影响。若干年后，缪斯尔在实验室里通过电视直播的方式为观众展示了计算机下棋的精妙。1962年，IBM公司的计算机一举击败了世界级西洋跳棋大师罗伯特·尼尔利（Robert Nealey）。一夜之间，IBM公司的股票飙升了15%。



图7.2 缪赛尔在IBM 7090计算机上下棋，摄于1956年2月24日

来源：IBM档案馆

计算机击败西洋跳棋大师之举令人印象深刻，但更令人诧异的是这套下棋程序还战胜了它的开发者缪赛尔。通常来说，基于规则的人工智能程序会受限于开发者的专业水平。然而，缪赛尔赋予了这套程序一种特殊的能力：如何学习。

许多人无法接受机器或计算机具备学习能力的说法。我们经常听到一些误解：“计算机程序无法比程序员更智能。”这种说法植根于我们对于计算机的陈旧观点，认为机器的自动化等同于去执行一系列预先设定的指令。但缪赛尔的机器学习设计，让计算机像人类一样具备汲取知识的能力，从自己的成败经历中学习。正如孩童能比父母的知识渊博，学生可能超越老师，运动员能战胜教练，计算机同样能比它的程序员更出色。

部分曾与计算机对弈的专家表示，在下棋的过程中，他们能感觉到软件对手下得专注、讲究谋略，甚至还有情绪。例如，加里·卡斯帕罗夫（Garry Kasparov）曾在国际象棋比赛中击败缪赛尔程序算法的伟大子孙——IBM深蓝超级电脑。1996年在《时代周刊》的采访中，卡斯帕罗夫说：“我能感觉到，对面坐的是一个拥有智能的家伙。”稍后他补充道：“虽然我认为自己发现了一些拥有智能的迹象，但那是一种怪异的智能，效率较低并且缺少韧性。我想电脑要想真正打败我，还得再等几年。”不过，卡斯帕罗夫错了。在第二年，也就是1997年，深蓝计算机彻底击败了他，捧走锦标赛冠军。

深蓝这种“怪异，效率较低的新型智能”，核心是数据的应用。机器学习即令人着迷，又让人沮丧，这都源自它的晦涩。所以，部分工程师抵触机器学习，就是因为他们从未完全弄明白人工智能是如何得出结论的——机器学习最具争议，也最让人难以理解的特点是其用于产生预判的非常复杂的内部数学模型。由于人类检测员无法通过软件代码发现系统是否正常运作，所以只能通过输入新的测试案例，来检测系统得出的预判是否正确。

无限状态空间

尽管机器人研究专家使用机器学习技术长达几十年，可这些机器人一直在高度结构化的环境当中运作。机器学习在棋类比赛中运作得非常好，这是因为棋盘是个有限的状态空间，所能产生的下棋方法数量是有限的。

塞缪尔的计算机下棋程序通过在数据库中的查找功能，使数据一一对应上棋盘的具体位置。每一个棋面布局都特征鲜明，定义清晰，因此容易进行数据存储。

下国际象棋比西洋跳棋的难度要高，因为每个棋局都有更多的未知可能性，这就是人工智能科学员所说的“高分支系数（higher branching factor）”。更复杂的情形还要数城市街道或是繁忙的高速公路，它们呈现的状态空间可能产生数不清的“下法”或“棋局”。在人工智能的研究中，能够为机器人提供无穷无尽新情形的环境被称为“无限状态空间（Infinite State Space）”。

无人驾驶汽车必须能够处理无限状态空间，因为它会不间断地遇到各种新情况。每一种情况都是全新的，不可能新建一个检索表来存储这些信息。不仅不可能把每一种新经验都归结为有限的可存储单位，而且即使能够生成，大量的存储数据还需要无比庞大的检索表才能迅速输出数据，这哪怕是现代最先进的计算机也无法做到。

多年来，无限状态空间一直是机器人专家把机器学习应用到非结构性的、充满变数的未知环境中去的障碍。直到21世纪初，机器学习才成功实现到无限空间状态的应用——现在已经出现新的算法，计算能力提升以及数量充足的训练数据发展成形。在2005年DARPA挑战赛中，斯坦福大学的工程师第一次尝试把机器学习技术应用在驾驶上。

特伦团队解决无限状态空间的方法其实很简单：把车外千变万化的客观世界简化成两种情况：可驾驶和不可驾驶。他们训练机器学习系统把车载雷达和摄像机收集到的原始实时视觉数据分成两类。为了教机器学习软

件识别可驾驶的地面，每个周末团队都会回到沙漠中，收集更多能反映沙漠地貌的视觉数据。当系统出错，他们就会及时调整，然后训练继续。

为了用采集来的数据给中层控制软件的概率网格图搭建视觉系统，特伦团队把数据流配上颜色代码。把机器学习软件认为可以驾驶的前方路面配上一种颜色，不能驾驶通过的路段数据配上另一种颜色。视频显示，当“史丹利”前进时，屏幕上会出现让人眼花缭乱的色彩漩涡，机器学习软件就这样把沙漠的无限状态空间简化成两种类型的。

“史丹利”在2005年挑战赛最终获胜，证明了计算机的视觉应用系统能通过机器学习在复杂多变的现实世界环境中应用。在机器学习软件的发展过程中，关键的推动因素之一是培训数据增多，数量充足。在无人驾驶的应用中，培训数据来源于车载硬件设备，曾经作为稀缺资源的培训数据，随着车载硬件设备在过去几年中的极大提升，也将变得不再稀缺。

现代工具箱

“重组创新（Recombinant Innovation）”指的是用创新的方式，把现行的几项技术组合起来的过程。尽管人们总觉得像无人驾驶这样的新兴技术（尤其是看起来非常复杂的技术）是孤僻的天才发明家用新奇的方式凭空创造的，但实际上它是重组创新的最佳案例。

重组创新是摩尔定律间接衍生而来的副产品，这条著名定律的内容是半导体材料的性能增长是呈指数型的，同时成本的下降也是呈指数型的。几十

年来，摩尔定律堪称指导半导体技术发展的真理。它的影响力已经延伸至其他类型的硬件，如内含计算机芯片的数码相机、数字电视和电动玩具。

自20世纪70年代以来，自动驾驶原型车逐渐得到高速发展，彰显了重组创新的力量和摩尔定律的影响。80年代，卡内基·梅隆大学的导航实验室研发出一辆名为“怪人（Codger）”的自动驾驶原型车，车型庞大如UPS全球特快的运输卡车一般。“怪人”的庞大体型是因为装载了多种昂贵的高端科技产品，包括大块头的彩色电视摄像机、GPS接收器、激光测距仪和几台安装了第三代太阳操作系统的计算机。“怪人”行驶在空旷的公路上，最高时速可达每小时20英里。该车行驶在城市街道上并不安全，在开阔的街道上完成导航需要10秒，若是在混乱的环境中需要20秒以上甚至更长……

让我们快进到2007年，此时已经出现了先进水平的自动汽车，前景一片光明。康奈尔车队为参加2007年DARPA挑战赛的车辆耗费巨资，用195,850美元购买激光雷达和传感器、GPS系统和摄像机；另花46,550美元购入若干台式电脑、手提电脑及外围设备。虽然在2007年装备一台无人驾驶汽车的费用要比20世纪80年代少许多，但购入的计算机和传感器的运行速度总是慢半拍，跟不上自动驾驶车速。2007年DARPA的赛后分析中，CMU车队的负责人，后来成为谷歌无人驾驶项目领导者的克里斯·厄姆森（Chris Urmson）悲哀地发现“现行可用的传感器无法满足城市驾驶的需求”……

再快进到现在，前景看起来更诱人了。车载中层控制软件能保证数据的供应，其价格也比2007年大幅下降。到写作本书的今天，装配一辆无人

驾驶汽车的硬件系统每年只需要5000美元，而且这个价钱在未来5~7年还会更低。现代硬件设备不仅价格亲民，而且体积小巧，适合安装在车体内部。雷达检测器的大小仅相当于一个冰球；GPS系统的接收器小到能轻易放置在汽车仪表盘内部；一台轻便的手提电脑处理能力远远超过1960年的微型货车般大小的大型机；一些雷达设备可以安装在汽车前灯隔壁。

今天的使能技术性能更好。就在厄姆森感慨“现行传感器性能不足”的7年后，2014年谷歌汽车的驾驶路程达到了70万英里。厄姆森说：“两年前，我们绝对应付不来城市街道的上千种复杂路况，而现在自动驾驶却可以处理得游刃有余。”2007年，无人驾驶还属于那种让人浮想联翩的“明日科技”。才过了几年，谷歌公司先进的无人驾驶原型车就成功在城市街道行驶超过70万英里。

谷歌汽车项目团队在短短几年时间内就设计和研发出一款功能完备的无人驾驶汽车，其成功也许可以看作以超前的意识走在行业前端的一个例证。此外，谷歌还享有其他几项优势，其中一项是资金。多年充裕的财政预算，足以使谷歌应付棘手的科研问题。

2007年，谷歌科研经费开支高达21亿美元，占公司年收入的12%。虽然不清楚21亿美元中有多少分配到无人驾驶的项目上，但我们可以做个对比，同年DARPA为其他所有团队分发了总共100万美元用于车子配置全套技术装备（含团队里学生购买比萨的钱）。

谷歌也有钱去支付精英人才的薪水。间接地，DARPA在系列挑战赛中的投资帮助谷歌公司突破瓶颈，加速发展。在DARPA挑战赛中崭露头角的精英随后成为谷歌公司无人驾驶计划“Chauffeur”的重要人才来源。塞巴斯

蒂安·特伦在挑战赛结束后不久就被谷歌公司聘用了。

在后来的采访中，特伦讲述了谷歌公司是如何在挑战赛的人才中挑选目标……然后再扩张至其他类型的人才。当DARPA挑战赛的精英人才都被挖掘并输送到谷歌任职以后，谷歌又吸引了世界各地多个领域的最顶尖的人才前来任职，研究领域包括机器学习、机器人、接口设计和激光技术。

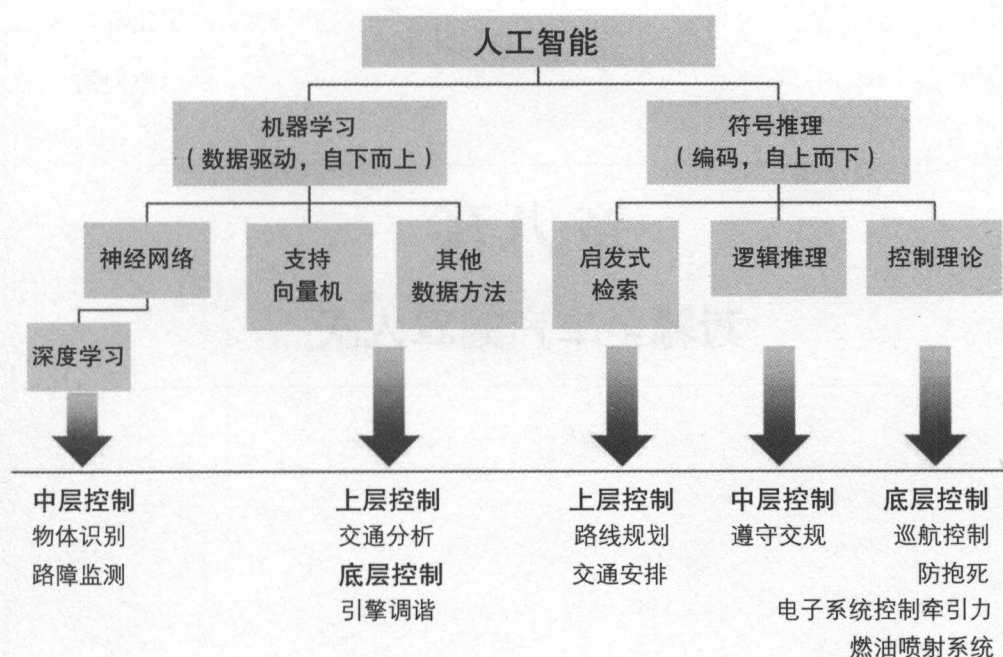
谷歌聘用的部分工程师最终为公司设计出，改装自丰田普锐斯（Prius）的第一代谷歌无人车。DARPA挑战赛的资深选手，安东尼·莱万多斯基（Anthony Levandowski），因为在伯克利大学上学时期就创造出全球第一部“无人摩托车”而扬名，他毕业后创立“510 Systems”创业公司。DARPA挑战赛结束后，“510 Systems”受聘于探索频道，给观众展示比萨外卖机器人。比萨外卖机器人的成功，引起了谷歌公司的注意。“510 Systems”的员工马吉思克（Majusiak）回忆称：“从那时候起，我们开始了与谷歌的合作……我们几乎承包了谷歌所有的硬件集成设备，谷歌只做软件。我们手上有汽车和底层控制系统，直接提供给谷歌使用。”2011年，5辆改装过的无人驾驶普锐斯问世，谷歌便整体收购了“510 Systems”。

资金和人才毫无疑问是大型重点科研先项目取得成功的关键因素。此外，还有第三种因素让谷歌公司的汽车在同行中脱颖而出：准备时间。机器学习软件，就像小孩子，需要时间去学习驾驶。历届DARPA挑战赛的参赛车辆是高校学生、教授和专业工程师团队长达12~18个月艰苦卓绝的心血结晶。由于DARPA挑战赛设立的规定，参赛公司没有充裕的科研时间，也不能私下在赛道上测试软件。为了确保比赛的公平，DARPA绝对禁止各个车队对车辆在街道、公路和模拟沙漠的酷热环境中进行赛前演练。车队只

能在无法知晓比赛会遇到什么障碍、路况的情况下调试软件。到了比赛当天，各参赛队伍纷纷把自家企业的声誉都赌在赛道上，向公众展示他们的车子。

事实上，为了保住“知名软件公司”“对数据有着独到洞察力”等良好印象，谷歌公司私底下进行了一次无人驾驶汽车的科研试验。因此，媒体和公众将无从得知谷歌早期的技术错误（如果曾经存在过的话）。到2011年谷歌的普锐斯车队面向公众的时候，车辆已经被“学习”得接近完美，表现无可指摘。

“充裕的科研预算、天才般的研发团队、预留的训练时间”，受益于这三大因素，人们看起来谷歌似乎轻而易举就创造出了优秀的无人驾驶车队。我们认为谷歌公司成功的原因里，还有一样和以上三大因素同等重要：绝佳的时机。事实上在过去的几年内，摩尔定律的无形而强大的影响力与日俱增，重组创新进入全盛时期。今天，无人驾驶汽车的发展终于步入正轨：智能软件来指路，数据支持有保障。提供数据来源的设备包括高速数码摄像机、高清数字地图、激光雷达和GPS定位系统。



汽车领域的应用

图7.3 无人驾驶领域的人工智能技术应用。大部分的机器人系统采用了各种技术的结合。物体识别技术被用来监测实时路障和识别交通秩序，部分是人工智能技术最大的挑战（如图中最左的一栏所示）

第八章

无人驾驶汽车全解读

Anatomy of a Driverless Car

无人驾驶汽车依靠不同种类的车载传感器回传的数据流来“听”和“看”，而车辆识别地理位置则是通过GPS定位系统和高清存储数字地图。让我们一起来深入了解一下，这些为汽车操作系统提供数据的硬件设施。

高清数字地图

人类通过识别鲜明的地标来熟悉新环境和认路，而无人驾驶汽车识别道路是靠GPS系统、视觉传感器和高清数字地图的指引。高清数字地图是详细、精确的模型，能提供某个区域最重要的地表特征信息。无人驾驶汽车用机器学习软件来处理实时交通路况，用内容详尽丰富、持续更新的高清数字地图来处理长时间的导航。

无人驾驶汽车通过高清数字地图上的GPS坐标来确定大致方位。然而，

GPS坐标与真实的距离存在几英尺的误差，这不能满足无人驾驶汽车的需求。于是，工程师需要想办法用其他技术手段来弥补GPS数据的低精确度，达到准确定位的效果。因此，早期的无人驾驶汽车比较注重数字地图的数据存储，其次才是实时GPS定位及其产生的数据。随着机器学习软件和视觉传感器（尤其是数字摄像机）性能的提高，汽车操作系统越来越精准地计算出当前位置，因为传感器能够精准捕捉周围环境的特征，并通过实时数据流形成可靠的视觉提示。

高清地图的精确度区别于标准数字地图。高清地图既能反映大型地理特征，像是高山湖泊，也能反应诸如树木和人行道的分布等微观地形细部。应用在无人驾驶汽车领域的高清地图，重点呈现的是道路或十字路口的表层静态数据，比如车道线、路口交叉点、施工地段以及道路标志。

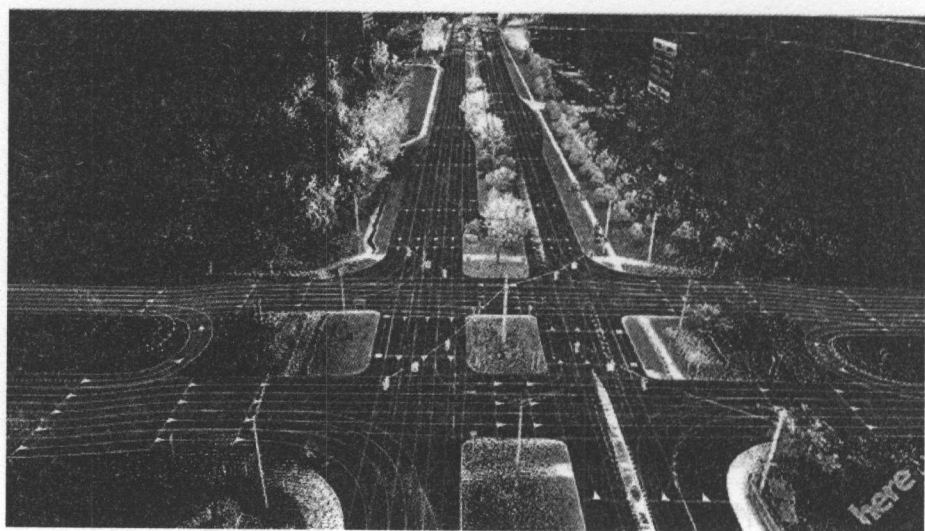


图8.1 覆盖着传感器数据的十字路口高清地图

来源：HERE地图

为人类肉眼设计的传统地图是二维图形成像，用静态标记表示某个地方的著名地标。而高清地图有着强大的封闭式后端，高清地图所呈现的某个区域的图像，其背后是存储着数百万个地形细部的庞大数据库，每个数据还包含着地理位置、大小、方位等其他相关信息。

人类的大脑都储存着一幅高清的本地地图。事实上，我们的大脑拥有惊人的“自动更新”“自动修正”能力，让任何软件工程师和制图员都嫉妒不已。更新一幅高清地图是一个艰苦的过程，需要载着摄像机和激光雷达走遍区域中的每个地方。这个过程我们将在本章后面详细讨论。

数码相机

数字地图存储静态数据，帮助识别汽车位置，而数码相机则类似人眼，在实时数据流中捕捉车外环境。随着数码相机技术的高速发展，拍照速度越来越快，图片也越来越清晰，机器人专家迫不及待地应用这些先进技术来提升中层控制软件的性能。

二十年前的1994年，苹果公司和柯达公司联合生产的QuickTake 100堪称当时最前沿的数码相机。当时的QuickTake以小巧便携和储存量大而著称，一次能拍摄八张 640×480 （30万像素）的彩页照片、重量是16盎司（450克左右）。今天，消费级数码相机平均每秒就能拍摄超过30张高分辨率的照片。

弄清数码相机的工作原理非常重要，因为数码相片的结构会直接反

馈至深度学习软件中。数码相机通过镜头以光子的形式采光。每个光子具有一定的能量。光子流通过相机镜头后，最终落在布满感光像素点的硅传感器上。感光像素点吸收光子的能量后转化成电能，储存为电荷。光线越强、光子的数量越多，电量就越高。光子击打在每个单元格的感光像素点上，随后被转化成计算机可读取的格式：一系列呈现每一个单独“图像元素”位置的数字合集，图像元素就是我们通常说的像素（pixel）。JPEGS、GIFs以及其他图像文件类型都是不同的储存途径。

数码相机借鉴了哺乳动物眼睛的一些概念。硅传感器在某种程度上与视网膜相似：两者的视觉数据都被拆分成几个小的视觉单位进行处理。视网膜上有数百万的生理感光细胞：视杆和视锥，吸收光子并把光能转化成神经信号，传输给大脑处理视觉信息。人眼内的视杆细胞和视锥细胞是不规则排列的，视网膜中部排列密集，边缘部分较为稀疏。而在数码相机里的硅传感器，每个像素在一定间距内呈矩形排列。100万像素的相机包含 1000×1000 个感光细胞，总共构成100万个像素。

一些适用于无人驾驶领域的专业数码相机功能，远远不止记录像素值。它并非从硅传感器的感光像素网格里直接输出光束的原始数字，先进的自动相机能自动分析图片数据。这样一来，图片的处理速度就能进一步加快了：相机先排除不相关的信息之后，再将数据输送给中层控制软件。

高级无人驾驶摄像机的处理得更到位：它首先处理图片包含的内容，把检测到的物体列成清单，并把结果制成表格。例如，这种相机描述的一个场景是：

- 1. 在左上方的角落有一个行人，以每秒1.23米的速度向左移动。
- 2. 最右边，消防栓，静止物体。
- 3. 左车道，一辆卡车以每秒5米的速度靠近。
- 4. 东南方向，不明物体，静止。

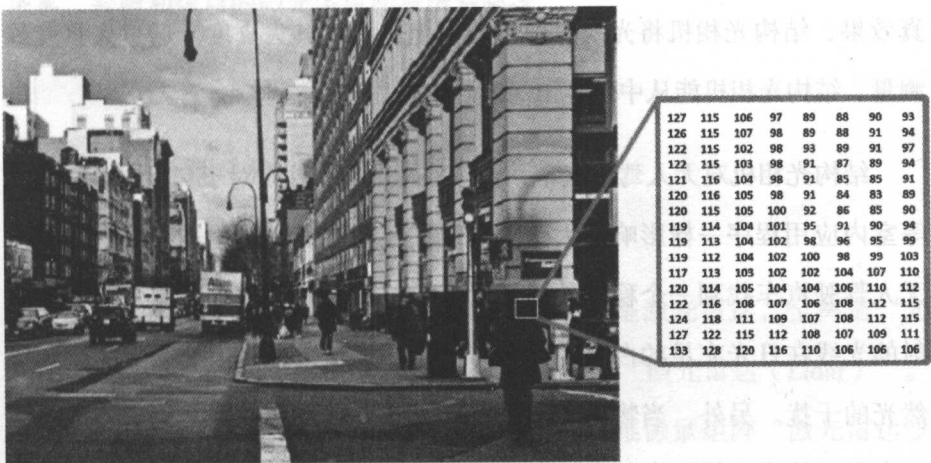


图8.2 图片拍摄于曼哈顿第十四大街，从第五大道往西看。左图为人眼所见，右图为相机所见。你能通过只观察数据，分辨出图片上的人和背景有何不同吗？
来源：维基百科

自然生物有两只或两只以上对等排列的眼睛，能够适应物体的深度感知，生物学家又把这称为“立体视觉（Stereo Vision）”。然而，数码相机是没有立体视觉的，这是限制数码相机在无人驾驶领域应用的最大问题之一。数码相机根据像素网格里的光线强度来捕捉信息，把三维数据世界转化成简洁的二维格式。但捕捉过程中，会遗漏一则深度感知的关键信息：物体与相机之间的距离。

科学家研究了多项不同技术来克服这固有的限制。解决方案一是在同一辆车内放置多台数码相机。在一辆无人驾驶汽车里，精心布置多台摄像

机来捕捉同一个镜头，每台机器的角度有细微的差别。这样能帮助计算机重建3D场景，对周围环境有更清楚的认识。

另一个可能方案是结构光相机（Structured-Light Camera）。使用拥有投影功能的数码相机，使图片数据拥有其深度信息。为了达到深度感知的仿真效果，结构光相机将光线投射到场景上并测量其失真度。通过失真度的测量，结构光相机能从中计算出深度。

结构光相机对无人驾驶汽车，就像Xbox Kinect外设对于交互式电子游戏等室内应用程序一样影响巨大，只是现在还不能确定结构光相机是否能为无人驾驶汽车找到一个稳定的“家”——结构光相机最大弱点之一是其投射的光线在日光充足的白天时段运行不畅，而且被投射出的光线会受到自然光的干扰。另外，当物体距离超过10米时，这种感知方案也不可行，这一点很可能成为结构光相机应用在快速移动的车辆上的致命缺陷。由于以上种种限制的掣肘，结构光相机最好的应用是在室内环境中，比如指引无人驾驶汽车出入停车场或车库，以及感应乘客的靠近车辆的移动。

数码相机持续呈现飞跃式发展，然而让人哭笑不得的是它还有一个低技术含量的弱点：污垢。即便是最好的自动数码相机，只要沾上一抹污泥，就不能正常工作。路边扬尘、沙土、鸟粪、虫子以及其他户外驾驶可能遇到的任一情况，都能妨碍最精密的数码相机和机器软件的正常工作。于是，我们能得出低技术含量的对应解决方案——给车载相机安装清洁剂，好比普通汽车的挡风玻璃上装有雨刷，也像人类的眼睛少不了眼泪的清洁和眼睑的保护。

许多针对数码相机弱点的解决方案，最终安装在汽车操作系统上。为

了确保车载相机的洁净干燥，每一辆无人驾驶汽车都要配备周期性自我检测的软件工具，它可对数码相机产生的数据质量进行定时自我检测。人工智能软件不断升级，为无人驾驶汽车提供的指引也越来越完备。将来的某一天，它能自动修正视觉数据错误，帮助无人驾驶汽车超越人类司机，在浓雾、暴雨和炫目的阳光下也能看清楚路况。

光检测与测距（激光雷达）

除了数码相机之外，另一类主流的传感器就是激光雷达，全称是“光检测与测距（light detection and ranging）”，亦称“激光雷达（Lidar）”。数码相机的工作原理是把三维视觉空间拆分为二维像素矩阵。激光雷达设备则不同，它向周围散射出强烈的脉冲激光，测算出光束反射回来所用的时间，然后根据激光测距原理计算出周围环境的三维数字模型。

和数码相机一样，激光雷达传感器的发展也遵循摩尔定律的轨迹，从20世纪60年代体积庞大、造价昂贵的固态设备转变至今天价格稳定的便携式设备。然而与数码相机不同的是，激光雷达的价格仍然高于人均消费水平。尽管每年的价格都在下降，2016年，Velodyne公司生产的重600克的16线激光雷达测距精确到厘米，售价高达8000美元。

几十年来激光雷达一直被勘测人员应用于获取详细的地形信息。相比之下，“安装这种设备，把激光光束射入移动车辆的环境”的想法，比它的整个发展过程晚得多。早期的现代自动驾驶机器人比数码相机出现得

还要早，为当时的无人驾驶汽车提供视觉数据的最经典传感器就是激光雷达。纵观三届DARPA挑战赛，激光雷达在获取车外路况方面扮演了极其重要的角色。堪称无人驾驶标杆的第一代谷歌无人驾驶汽车，改装过的普锐斯使用的就是激光雷达。

激光雷达能绘制出高精度的3D模型，并附有准确的深度知觉，这就决定了激光雷达在无人驾驶汽车领域不能撼动的重要地位。激光以大约每秒10亿英尺的速度传播，所以激光雷达传感器能利用若干台千兆赫兹（GHz）的微处理器进行深度测量，分辨率可以精确到厘米。

激光是理想的测量工具。蜡烛或白炽灯的光向四面发散，而激光能够呈直线定向发射至很远的距离。激光是平行光，它不会像手电筒的光一样发散，也无论在传播过程中是否遇上物体阻挡，物体的距离在一步之遥还是百米之外，它的传播均能保持准直。为了构建完整3D数字模型，激光雷达需要高速旋转，不断向四周扫描。其工作原理就是用一组同样旋转的镜子使激光改变方向。

假设有一间堆满隐形物体房间，我们可以用一罐红色涂料喷涂在所有的隐形物品上，直到它们完全可视。如果只有你一个人，那么需要很长时间才能让所有物品都现形，如果好几个人同时手持红色喷漆，这些物体很快就能被颜色覆盖，进入视野。激光雷达的工作原理也是这样。所以激光雷达需要同时发射多组激光，对周围环境平行地进行扫描和测量，而激光光束越多，形成的数字场景模型分辨率越高、速度越快。

一台无人驾驶汽车的数据由激光雷达产生，传输给软件，再由软件把信息整理到名叫“点云”的数字模型里。如果激光束直射入高空，所见模型为

一片空白，因为发射路径中没有能够反射激光的物体。与此相反，如果激光束射入车水马龙的街道，所形成的“点云”将获取各种有趣的详细信息。

雷达的激光束以某种特定的模式朝外射去，形成数字“点云”的过程其实与全息照片呈现在底片上的过程相类似。随着旋转镜把激光调整为一系列平行的光线，迅速地识别了车辆行驶前方的道路。激光雷达构建出来的数字点云是由大量结构细密的扫描线组成，数字模型中的每一行与旋转镜上的每一根扫描线一一对应。图8.3呈现的是激光雷达所产生的三维“点云”，在图中，你还能发现从水平扫描产生的圈状扫描线。

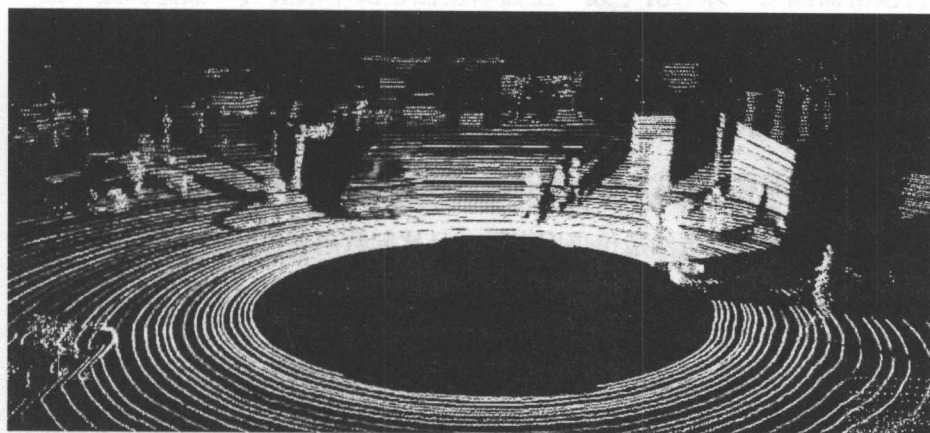


图8.3 这张三维“点云”数据由车载激光雷达拍摄。当时车正经过一个喧闹的停车场，在图上可以看到一些车辆和行人

来源：Velodyne

若是一个轻易下结论的观察者，可能会认为激光雷达“点云”与数码相机几乎是一回事。事实上，激光雷达与数码相机在许多方面存在很大的不同。其中一个重大区别是，激光雷达传感器无法捕捉色彩信息。激光雷

达所产生的扫描结果，就像拍摄于某个暴风雪过后的照片一样，呈现的结果将是幽灵般的同一颜色。实际上，在软件在处理“点云”信息时，会手动给近处物体涂上蓝色，远处物体用红色，黑色则表示没有任何物体，或者激光没有反射返回。

激光雷达点云与数码摄相的第二点不同是成像时间。激光雷达不停旋转，持续更新所产生的数字模型。一方面说来，“点云”的不断更新是件好事；另一方面，这使得“点云”与数码相机的瞬间“咔嚓”按下快门一点也不相似。激光雷达传感器速度慢，虽然能高效绘制出地形轮廓图或慢慢挪移的交通堵塞数据图，但是不能及时提供某些紧急交通状况下，计算机需要处理的瞬时影像数据。

今天的无人驾驶汽车既使用数码相机，也使用激光雷达。在几十年前人工智能发展滞后的年代，激光雷达就已经是必不可少的视觉传感器了。如今的激光雷达传感器虽然造价仍旧高昂，且比起数码相机要慢半拍，然而它产生的点云却能为车辆规划绝大多数环境中的路线，并指引其完成驾驶。

长期阻碍数码相机成为机器视觉传感器的“瓶颈”是三维感知。而且解压、处理大量像素需要强大的计算能力支持，在实时操作中计算能力却往往跟不上——这是数码相机在无人驾驶应用上的严重缺陷。不过在过去几年的高速发展后，数码相机也终于成为无人驾驶领域的应用工具之一。随着微处理器的速度不断提升，数码相机和软件处理数字图像的能力也与日俱增，也许不久的将来数码相机机会取代激光雷达，摘取视觉传感器的冠冕。

部分专家认同这种看法，在2015年10月特斯拉公司的一场新闻发布会上，埃隆·马斯克对该公司的未来无人驾驶汽车的技术发展评价如下：

我认为人们无须使用激光雷达，无源光（passive optical）外加前置雷达就能做到这一切。事实上，即使没有激光雷达也能彻底解决问题。我不看好激光雷达，它在这种情况下不起作用。

无线电检测与测距（雷达）

除了激光雷达和数码相机，无人驾驶汽车还使用普通雷达（Radar）来识别周围环境。数码相机在像素网格中捕捉场景，激光雷达传感器则类似一罐“数码喷漆”，让原本无形的物体披上“数字化外衣”。然而雷达传感器的工作方式就像往池塘里扔下一块小石子，追随水面漾起的圈圈涟漪跳跃的方向。

雷达原本应用服务于军事领域，在二战期间，雷达发射塔建在海滩、田地里，用于侦测敌军的飞机、船只和来犯的导弹踪迹。战后，美国把雷达用于航空交通管理，监控和确认商业航班的飞行航线。许多从高速公路警察手里接到过超速罚单的人，也能直观感受到了雷达技术的威力。

作为摩尔定律的又一力证，雷达传感器体积轻巧、性能稳定，完全可以在快速行进的车顶上正常工作。在现代普通轿车中，雷达传感器一般应用在定速巡航技术方面：通过内置雷达装置感应到车辆周围的其他车辆的速度和位置，定速巡航系统据此调控刹车和油门。雷达传感器另一个常见的应用是驾驶辅助系统：当有车辆靠近司机的视线盲区时，系统会发出警示。

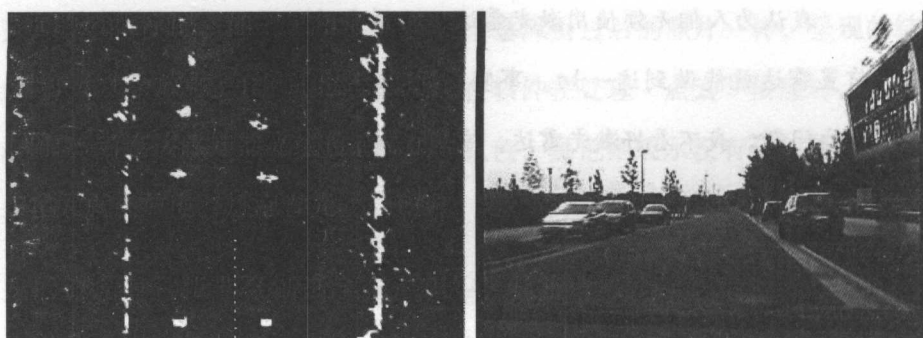


图8.4 前置雷达的原始目标密度图（左）和车窗前对应的景象（右）。大型静态物体（如停泊的汽车、建筑栏杆、路灯等）都反映在图片上。本图片由基于网格划分的24Ghz雷达传感器生成

来源：Smartmicro 3DHD

雷达传感器由发射器和接收器组成，发射器负责向外发射电磁波，而接收器随时准备接收回波。雷达传感器利用电磁波的反射来检测周围环境中物体的出现，先是发射一系列向外辐射的电磁波。如果电磁波在传播路径中没有遇到障碍物，它们会继续呈环状向外扩张，直到最后消失在远处。如果电磁波在传播路径中遇到障碍物，则会发生衍射，绕开障碍物，改变传播方向。由于电磁波的速度等同于光速，因此整个传播过程非常快。

由于回波的能量比起发射出去的电磁波能量大大衰减，接收器使用了电波放大的技术确保不遗漏那耳语般微弱的回声。为了防止传感器意外接收到附近雷达发射出的电波，每个雷达传感器的电磁波都会拥有独特的签名——编码脉冲技术“啁啾（Chirp）”。回波形状和回波时间都隐含着障碍物的形状、构成材料等信息。一些雷达传感器通过分析回波频率的变化，计算出障碍物移动的方向。

电磁波的波长是指相邻两个波峰之间的距离，不同的雷达传感器波

长不同。电波之间间隔越大——波长越长——传输得越远。长波在传播时更容易绕过环境中的小物品，因此长波监测到的环境图像更粗略。短程雷达传感器向外发射微波，微波能监测到许多细小物品，可以小到猫咪般体型，轻便如自行车一般的物体。另外，电磁波在高导电性的物体表面反射效果最好——例如平整、光滑的表面，闪亮的自行车表面和潮湿的路面。而非导电性物体，比如多孔塑料或木头制品，在雷达看来几乎是“透明的”，很难监测。幸运的是，绝大部分汽车，即便是那些安装了不少塑料配件的汽车，内部也有足够的金属零件，足以使雷达传感器监测到它们。

一台雷达传感器只能在特定的、狭窄的范围内进行识别。所以大部分传感器都是成排安装，传感范围稍有重叠。无人驾驶汽车有个典型的现象，就是三台雷达传感器必须相邻安装，以便提供180度视觉覆盖效果。

对于自动驾驶来说，雷达传感器最大的好处是可以穿透浓雾、暴雨、尘土、扬沙，甚至是强烈的顶灯灯光，这一点不同于数码相机。雷达传感器的另一个优点是能轻易穿透非导电性物体和薄型材料，因此不容易受到高速公路上空飞舞的塑料袋或风滚草的干扰。电磁波更能识别大型的物体，因此司机担忧的巨型障碍物容易引起它的注意。相反地，雷达传感器最大的缺点就是其分辨率相对较低。

雷达传感器还有一个优点，它不仅能检测到物体的位置，还能根据多普勒效应监测到物体的速度。多普勒效应是以19世纪的奥地利物理学家克里斯蒂安·多普勒（Christian Doppler）的名字命名的最常见的物理现象之一。一个人站在高速公路旁，听到疾驰而过的车辆发出轰鸣声，声音的音调随着车辆远去而降低。音调降低是因为高速运转中的汽车引擎产生的声

波，当车辆驶近时，声波受到压缩，所以音调升高，当汽车呼啸而过，声波延长，音调降低。

雷达传感器利用多普勒效应得出物体的速度。通过记录电磁波向外发射和反射回波的频率变化，判断移动中的物体是正在靠近还是渐渐远离，传感器还能计算出物体移动的速度。速度信息能帮助我们分辨该物体究竟是什么。比方说，在公路上每小时30英里的速度飞驰的不太可能是行人。

在无人驾驶汽车中，雷达监测与视觉传感器在周边环境监测方面的作用相辅相成。通过感应周围物体的大小、疏密、速度和方向，雷达传感器传输的许多信息能与数码相机、三维“点云”生成的图片相互参照。路边静止不动的小东西是只猫还是个硬纸皮盒子。注意，一个大型的“金属盒状物”快要接触我们的尾部了！

相比以前，今天的雷达监测智能化程度更高。二战期间的雷达在绿色屏幕上呈现原始模拟回声信号，外加一根不停旋转的天线表示目前的扫描方向。现代雷达的传感器类似于先进的数码相机的工作流程——能够预先对原始信息进行处理，列出包含尺寸、地点和速度等内容的目标清单。经过处理的信息更简洁，占用更少的存储空间，对于无人驾驶系统来说更易懂。

为避免监测的情况过于琐碎，雷达还能尝试着把公路路面上的不必要细节从目标清单中去除，但是沥青路面也和其他物体一样会反射电波，形成干扰。因此，部分雷达传感器会自动去除扫描结果中静态物体的数据记录。虽然从目标清单上屏蔽部分物体能提高效率，同时也带来一定风险。雷达可能把某些巨型、静止的物体，比如停在桥下的一辆熄火的汽车，误认为是桥墩的一部分。

超声波传感器（声呐）

如果把激光雷达和数码相机比作一个人的眼睛，超声波传感器就是人的耳朵。超声传感器可谓雷达的近亲——像雷达一样，发射波形信号，检测回波——不过超声波传感器发射的是超声波，不同于雷达的电磁波。

超声波传感器根据物体反射回来的时间、频率、声波形状来监测位置和速度。超声波传感器一般分为两个部分：发送器和接收感应器。发送器产生振动频率高于20kHz的机械波，高于人耳能听到的频率范围；接收感应器感应到声波的回声并对其进行处理。

超声波传感器有着许多与激光雷达、普通雷达相似的优缺点。就优点而言，它与雷达传感器一样都能穿透浓雾、尘土，能在刺眼的阳光下看清物体。声波的传输速度比电磁波慢上许多，这决定了它能以高分辨率来看清细小物体。同时由于声波的能量随着风和距离快速衰减，它只能近距离监测物体。因此，超声波常作为雷达实际应用中的补充，比如停车等近距离的精确检测。

全球定位系统（GPS）

虽然数字地图和各式雷达传感器在无人驾驶方面扮演着极其重要的角色，可是除了这两者意外，还有一项更为关键的技术——GPS全球定位系

统。这个系统负责统筹和整合信息，并在高清数字地图上为汽车提供最终的精确定位。

GPS是一项已经面世长达数十年的技术，最初用于军事领域；发展遵循摩尔定律，后来发展成为稳定可靠的、低成本的消费应用品。几十年前，GPS的接收器像冰箱一样庞大，而今天已经变成了小小一块芯片，嵌入手机、相机、平板电脑和汽车里。

GPS接收器能接受太空卫星发回的信号，堪称尖端科技发展创下的奇迹。车内或手机里的GPS接收器通过聆听太空卫星发出的蜂鸣信号声，确定你所处地点的经纬度。每一个卫星遵循预先设定的轨道，能持续发出稳定的电脉冲流，精确到以秒为单位。

太空中总共有24颗卫星为GPS系统提供信号，但GPS接收器只需要接收到4颗卫星就能计算出地球上的具体定位，而其余卫星的信号则更能帮助设备进一步做测量。每颗卫星发射自己专属的信号，GPS接收器把这些信号汇集后形成一股特有的信号，发给原有的卫星。工作中，卫星信号源源不断地输入，GPS接收器必须仔细分辨聆听。通过计算信号响应之间的时间间隔，GPS接收器应用“三角定位”的数学方法计算出准确的位置。如果两颗卫星的信号同时到达，GPS接收器把这种情况判断为位置恰好处于距离中点的平分面上。

在常见的驾驶环境中，GPS接收器能计算距离为4米，约13英尺。如果接收器提供的定位信息精确无误且持续更新，研发无人驾驶汽车的难度就可以大大降低。遗憾的是，由于大气环境中的云、雨等造成干扰，卫星信号可能遭遇屏蔽或延迟，导致计算结果产生偏差。

都市驾驶的另一个严重问题是超声波的反射脉冲。如果你曾经在纽约的曼哈顿使用GPS设备，你一定感受过卫星脉冲信号遇上高层建筑产生反弹的情况——GPS系统出现混乱失控，节奏凌乱，每隔几分钟就产生新的定位。事实上是，从卫星上发出的部分脉冲信号，遇上了摩天大楼产生反弹，给接收器造成假象，误认为它们的抵达时间略有不同。就连最好的GPS系统也能被城市峡谷效应（Urban Canyon Effect）误导。

内置的耳朵（IMU惯性测量单元）

GPS的失误会造成灾难性的后果。卫星感知的缺失这个潜在的致命难题的解决方案是另一种军事技术的产物——惯性测量单元。这种装置有两种重要功能：它能弥补GPS的不准确计算；它还是无人驾驶汽车的“内置耳朵”，能够感应，也就是决定该走哪条路。

惯性测量单元是一种多用途多功能的设备，它配有加速感应和定位感应，能记录汽车行驶的轨迹，从而也能看出诸如左右轮胎高度是否一致等问题。现代的惯性测量单元是一系列复杂设备的集合，包括里程计、加速计、陀螺仪和指南针。所产生的结果是各个设备数据经由复杂的预估算法解析后综合得出的。

惯性测量单元是个独特的传感器，其传感范围局限于汽车主体之内。人类本身拥有一整套类似的感应系统，称为本体感觉（Proprioceptive Senses）。本体感觉，不同于我们外在的视觉和听觉，是对我们体内进行观

察。平衡感就是一种本体感应。当你闭上眼睛坐在一辆驶离车站的列车上，你能感觉到加速，不用亲眼看见也正在向前行驶。这是另一种本体感觉。

惯性测量单元使用了一种老式的导航技术——推测航行法，在GPS数据和修正其不准确率之间保持准确的汽车定位。几个世纪之前，水手在无边无际的大海中航行是依靠参照星星的位置判断方向。然而，当暴风雨来临时，星星被乌云遮蔽，问题就棘手了。推测航行法是通过水手上一次看到星星的时候来推算船走了多远，从而推出船目前的方位。通过测量相对地理位置，而不是绝对位置，水手能确保船只大多数情况下不偏离航线，直到天空放晴，星星再一次出现来指引他们。

推测航行法还能运用在以下情形中：多云天气来临时，当船照常向前航行，水手会往船后扔下一捆绳索，上面间隔地系着常见的绳结。他们飞快地数算绳结浸入海面的数目，这样就能算出船速。甚至在今天，人们依然使用绳结法来计算船前进的速度。一旦水手得知船前进的速度，用指南针测出航行方向，他们也能算出从上一个已知的航行地点（我们常说的定点）到现在航行了有多远。在无人驾驶领域，当车子驶入隧道或是穿过紧凑布局的“城市峡谷”，接受不到卫星信号时，惯性测量单元也会使用类似的方案来解决。

今天我们不用计算绳结往车后飞去的速率，汽车上的里程计在计算从上一个地方至今的里程，是通过统计车轮旋转的圈数。尽管车轮圈数是相对精确的机械数据，车程的累计仍然有些不确定因素。轮胎气压变化或汽车变换车道时，轮胎可能打滑。经过高速公路的环形路段时，里程计最后可能会统计出不同的车轮圈数，这取决于汽车行驶在内车道还是外车道，

由于两车道的距离差异，统计的里程数可相差数十米。

由于简易的里程计的读数并非完全准确，惯性测量单元通过启用加速感应器来解决该问题。当汽车匀速行驶，加速度记录数为零，这与我们的直觉相悖。只有当汽车增速，减速或突然转向，加速度才有变化。若GPS失灵，惯性测量单元则结合加速计与里程计的数据进行计算。然而加速计本身不提供汽车行驶的方向。这时指南针开始发挥作用。惯性测量单元加上GPS系统和指南针构成一个强大、万无一失的组合。而且，惯性测量单元的功能不仅仅是充当GPS失灵的替代品，还能为无人驾驶汽车提供平衡感。

机器人专家把机器人的空间方位感称为“位姿（Pose）”，一辆汽车的“位姿”是车前端指向的方向和车身倾斜的角度。要测量“位姿”，我们需要在惯性测量单元的组合中再增加一个感应器：陀螺仪。陀螺仪（分为机械陀螺仪和光学陀螺仪两种）是个高速回转体，用于测量位姿。惯性测量单元需要获取三种信息才能测量和记录汽车的物理空间方位：车辆前进的方向、车前端的俯仰角度和倾斜侧角度。此时，古老的航海技术再一次在现代惯性测量单元身上留下了烙印。古代造船工匠和现代的航天工程师不约而同地将这三种测量车或船的维度称作：偏航角（车辆左/右转的角度）、仰俯角（船只或车辆发出的噪声高低）、翻滚角（从一边到另一边的倾斜度）。

即使GPS系统运行良好，确定车辆的“位姿”仍是一项重要的安全机制。假设汽车在冰封的路面上开始打滑，它的偏航角从零度猛增至360度（原地打圈）；当它走下坡路时两个车轮打滑的程度不同，则仰俯角就会开始前移；如果这种状况持续发生，那么当两个车轮脱离地面，车子开始

滚动。

惯性测量单元实时测量汽车的偏航角、仰俯角和翻滚角，并把数据输入软件，好让驾驶系统及时调整：打滑时踩下刹车，可能翻车时发出危险警报。今天大多数的普通汽车都安装了用于记录车辆行踪的内置惯性测量设备，使车轮上的减震器更加平稳。无人驾驶汽车，同样离不开惯性测量单元。

现代惯性测量单元是摩尔定律的又一例证。这个设备始于二战期间，比硅晶片出现得还要早，是一种结构复杂的机械设备，最初被用于计算火箭发射的最佳弹道。

最初一款高档、制作精良的设备，只能被应用于商业船只和潜水艇，其造价高达100万美元。但服从于摩尔定律，其成本在不断下降。20世纪80年代，惯性测量技术受到软件和微传感器问世的影响而转型。以微传感器为代表的微机电系统（micro electro-mechanical systems）大大改变了惯性测量技术，使它从昂贵、专用于太空飞行和军事领域的设备转变为相对小型和平价的导航系统单元。如今，每部手机里都能嵌入简单的惯性测量单元。大多数智能手机能够通过指南针来分辨方位。如果你在长途飞行中，可能会看到有的乘客在玩游戏时，不时轻轻晃动或左右摇摆手中的iPad，这就是惯性测量设备的具体应用。

除了高价以外，惯性测量技术最大弱点就是不能脱离GPS长时间工作，否则会产生偏移。各个微传感器所产生的微小偏差累积起来就成了大问题。缺乏准确的卫星数据指导，就像古时候的船只在航行中遭遇好几周的多云天气，即使再精确的惯性测量设备也不能准确测定航线。

线控技术（Drive by Wire）

数码相机、激光雷达、普通雷达、超声波传感器和惯性测量单元等众多传感装置为无人驾驶提供稳定的数据流。数据流融合后，汽车的操作系统顺利进行数据处理，奇迹就这么发生了。正如我们在前几章提到的，汽车操作系统使用好几种人工智能技术来快速做出指令。最后一步就是把这些指令转换成实际操作，像是转动方向盘、踩下刹车或油门。

在以前，工程师把普通汽车改装成为无人驾驶汽车，是通过装配特制的机械“线控”装置，目的是取代人类司机的手和脚。这些新装置叫作线性致动器（Actuators），通过它们可以控制方向盘，踩下刹车。生产出准确、稳定的机械驱动器是工程学的难题，其过程的复杂不亚于创造出人工智能来指引汽车。

汽车的子系统在过去二十年中自动化程度越来越高，研发无人驾驶的人造“肌肉”工作难度相应降低，液压机械系统逐渐被计算机系统所取代。大部分现代汽车配备好几台计算机指引功能的子系统（底层控制系统），系统内置了微处理器，处理上百万行代码不在话下。今天的机器人专家不再使用特制的机械“脚”去踩油门，只需在汽车的电子系统上稍做改动即可。

软件好比掌控机器的精灵。无人驾驶汽车使用多个电子通信系统，包括操作系统、高/中/低水平控制等来解读指令。今天平均每辆汽车都安装好几个子系统，例如引擎控制单元（ECU）、ABS制动防抱死系统、自动变速

箱控制单元（TCU）。这几个系统通过“总线（Bus）”互联互通。

在计算机术语里，总线是指一个交流频道，把数据从计算机内部的一个部件传输给下一个。街道上的巴士和计算机数据“总线”有着相似的语源学渊源，它们都来自于拉丁文Ominibus，意思是“公共所有的”。就是城市里的巴士把乘客运载到各处，无人驾驶汽车里的数据也是这样被总线传输到各个子系统，也像通用串行总线（universal serial bus，USB）连接计算机键盘、鼠标和打印机。

许多汽车今天使用的是控制器局域网总线协议（controller area network bus protocol，简称CAN总线协议），以1Mbps的速度往返传输数据。CAN总线协议是一种“点对点”的协议，由ISO11898和ISO11519两个国际标准来管理。作为公开的国际标准，意味着任何设备都能接入CAN总线，也能“理解”协议，该协议适用于各个车型。

一般情况下，汽车公司不会公开宣传控制协议的内容。然而，他们会与其他制造商共享那些接入车内控制系统的新设备所需要的信息。或者，他们会给某些制造商分享网络协议的详细内容，比如说，那些把底盘售卖给房车生产公司的汽车零件制造商。

汽车控制器局域网也和其他的网络一样，最关键的要数带宽（bandwidth）和网络稳定性。带宽是指数据在网络中传输的最大速率，通常以每秒多少bits为单位来计算（即bps）。当被应用在无线网络中时，有形的网线被无线电波或是不同的波段所取代。网络带宽是由微处理器对电脉冲编码、解码的速度决定的，还有就是总线中能同时容纳多少平行波段。

带宽在任何网络传输中都非常重要，但对行驶中的汽车而言，很多时候，数据的传输速度则更为关键。大多数无人驾驶的数据传输系统使用一系列通用的代码来呈现特定的动作，这样能节省时间。例如，无人驾驶汽车下指令“刹车”，该车的软件系统早已准备好对应的两位数字表示“刹车”，这样就能马上抵达子系统。因为两位数是小而高效的意义单位，系统只需要16微秒进行转换和接收信息。16微秒的响应时间对于无人驾驶汽车来说非常合理。它甚至比一次眨眼的瞬间快上1000倍（一次眨眼的时长是100~400毫秒）。虽然CAN总线能快速传输微小的数据单位，但如果汽车CAN总线需要处理从各个传感器中涌流而出的数据流时，“带宽”就会迎来挑战。

当汽车传感器实时输出大量的数据流，系统速度会下降。如果某网络的传输速度是1Mbps，把1MB的图片从相机传输到中层控制系统，实际上会需要漫长的8秒时间（1Mbps是1兆比特/每秒，1MB是1兆字节，1字节是8比特）。假设现在加重网络负担，数据从汽车的其他传感器中流入，1Mbps的带宽就显然无法满足无人驾驶的需求。承受着实时视觉数据的重负，汽车CAN总线的响应时间会变得慢吞吞的，这在实际驾驶中是不能接受的。

展望未来，汽车制造商需要建立一套无人驾驶汽车专用的稳定、透明的通信标准，这样才能处理传感器传输过来的数据流，同时防止数据外泄。换言之，无人驾驶汽车需要一套高带宽的总线。有两种基本方式适用于解决任何网络中的通信瓶颈：第一，增加可用的电子线路或波段数目，平行发送数据流；第二，使用压缩算法，大量整合数据，形成更高效的单元。而且，数据压缩的方法还可以直接在传感器上进行。例如，一些车载相机包含实时图片分析软件，经分析和压缩后，把图片相关信息一并发送。

除了带宽，稳定性是车载网络另一个重要特性。无人驾驶需要像下载的音乐一样顺畅，还要比金融交易的过程更为安全，所以数据传输的稳定性体现在以下几个方面：

其一是防御的能力。当出现不请自来的恶意制造出破坏网络稳定性的第三方设备时，CAN总线顿时变成一个战场，这让人联想起晚餐约会时擅自闯入了一位宾客，他打断了餐桌上的对话。这种恶意的装置不仅仅是妨碍，一旦入侵车载网络，甚至能劫持整辆无人驾驶汽车。

其二是网络容错性（tolerance to errors）和消除网络噪音的能力。无人驾驶汽车需要一套足以媲美航空电子设备的弹性高效的纠错协议。

假设无人驾驶汽车的软件发出指令“油门增加1%”，却被燃油注入系统误解为“油门增加100%”，那么公路将会上演一场血腥的屠杀灾难。为了防止系统间数据交流出现致命的错误，纠错协议提供监督功能，就像聘请了一位冷静笃定的校对，二次检查信息传输的内容。无人驾驶汽车CAN总线上的各个子系统需要互相信任。良好的通信协议为子系统核实所收到的信息与传感器发出的内容一致。

鉴于人类司机在车内驾驶时应对恶意攻击能力的严重不足——无论是自动驾驶汽车还是普通汽车——我们都希望汽车制造商对车内通信协议进行加密保护。遗憾的是，这种安全防范思维还没有稳固地植根在汽车行业里。也许是因为部分车主热衷于摆弄汽车的引擎，所以如今入侵一辆汽车并不太困难。

其他设备

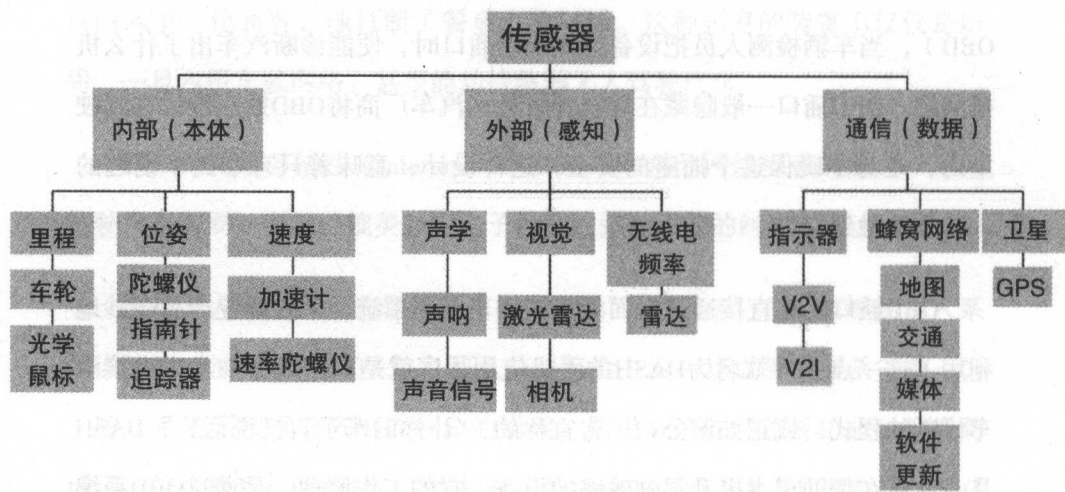
另外，汽车现存的某些明显“弱点”，实际上是有意这样设计的。大部分的汽车都有个实体连接器，叫作车载诊断系统（on-board diagnostics, OBD），当车辆检测人员把设备插入OBD插口时，便能诊断汽车出了什么机械问题。OBD插口一般隐藏在驾驶杆附近，汽车厂商将OBD插座设计在驾驶室内，是为了确保这个插座的安全，这样设计，意味着只有拿到车钥匙的人才能接触到它。

OBD插口能够直接通过后面板连接汽车操作系统，某些商业产品巧妙地利用了这一点。一款名为DASH的手机应用程序就是利用蓝牙技术来连接汽车诊断的程式，或正如该公司广告宣称的：“让你的车子自己说话！”DASH是一款意在帮助车主提升驾驶效率的设备。它的工作原理，是通过OBD系统“窃听”车辆的数据，甚至可以记录司机踩了多少次油门和刹车。

DASH同时也整合数据，公布某些驾驶数据，例如：在哪个路段司机常常急刹车或急转弯。由于DASH知道你所有的驾驶习惯，所以老板和市场推广人员都会对你的汽车很感兴趣，因为它就像是你的网页浏览喜好。因此，DASH这类设备面临着个人隐私保护的问题。

我们的技术是否达到无人驾驶汽车的水平了呢？从硬件传感器所能提供数据来看，答案是肯定的。今天传感器套件的质量和价格绝对能满足无人驾驶的需求。事实上，随着摩尔定律持续生效，传感器会继续提升速

度、降低价格；呈指数型增长，每过几个月传感器性能就会翻番，价格减半。现在让我们把注意力转移到软件上，深入研究那充满期待的“最后一块拼图”——深度学习技术。作为软件控制系统冠冕上的明珠，它为机器人感知与响应提供指导。



汽车应用

速度控制	物体识别	全球定位
线路控制	路障监测	与道路基础设施交互
GPS修正	其他汽车感应	发送/接收最新交通信息
运行状况监视		发送/接收地图数据
		信息娱乐
		软件版本更新

图8.5 无人驾驶汽车中应用的重要传感器，绝大部分的自动驾驶汽车都将使用以上多种传感器组合

第九章

深度学习，无人驾驶的最后一块拼图

**Deep Learning
The Final Piece
of the Puzzle**

人类识别物体，是根据物体本身的特征，例如从标志性的特征认出某位朋友，或是根据某个颜色鲜艳的标签，让你一眼就从行李传送带上堆积如山的行李中认出自己的行李箱。现在软件也能根据物体显著的特性来学习识别某个物体。

人类有持久稳定的记忆，能够认出在不同情况下见过的物体或场景，深度学习最终也能赋予机器这种神奇的能力。哪怕数码照片在机器不熟悉的环境和不同层次的灯光下拍摄的照片，深度学习也能识别照片中的物体。事实上，深度学习软件通过不同的应用，已经达到与人类水平相当的识别能力。我们也许不久将会看到莫拉维克悖论解决方案的诞生，因为机器人专家和计算机科学家已经找到新的方式，把深度学习应用到机器人感知与响应领域。

2012年以来，深度学习发展日渐成熟，现在已经能够应用到多个领域。在无人驾驶汽车方面，深度学习能够为车辆提供视觉能力，并提升语言识别

软件的语言理解能力。2016年，深度学习软件在一次备受瞩目的活动中显示了其多才多艺——该活动就是谷歌公司的AlphaGo围棋项目——人工智能软件击败了世界上最顶尖的棋手——要知道围棋可是被公认的比国际象棋比赛更有挑战性。为了鼓励第三方开发商使用其软件工具来开发智能应用程序，谷歌、微软和Facebook分别推出了各自的开源深度学习发展平台。

与其他深度学习软件类型一样，深度学习软件需要极高的计算能力和大量的训练数据。以高速计算机和传感器为代表的使能技术发展成熟，并不是推动深度学习软件得到普及的唯一因素。政府因素，也是促使深度学习高速发展的巨大动力。

数十年来人工神经网络的发展步伐缓慢——但这项技术决定着深度学习软件的基础——因此一系列的意识形态都在高校计算机系的研究中停滞不前。自从1950年作为一项正式的研究领域出现以来，人工智能领域就是一项意识形态的战争。由于没有人真正明白大脑是怎么工作的，几个世纪以来，哲学家围绕人类思想的未解之谜展开了激烈的学术斗争。大脑和心智之间的不同在哪里？智能是什么？知识又是什么？现代学术界致力于研发人造大脑的科学家，发现自己和历史上的哲学家一样无法解答上述难题。

在人类知识的正式研究过程中，研究学者总是会倾向于与某种思想流派，并常常为了捍卫该流派而穷尽毕生之力。虽然有别于思想流派之间争论，但人工知识的研究实际上也遵循着这种规律。对于高校学术界以外的人来说，关于众多不同人工智能方法孰优孰劣的学术辩论就像是一场计算机抽象化理论的斗争。然而对于这些学术研究学者而言，这些辩论一点儿也不抽象，实质上这是关于稀缺资源的战争。

年轻的人工智能研究员往往进步神速，他们所认同的思想观点决定了他们职业生涯的发展轨迹。教授和研究人员判断一项人工智能研究方式是否合理，是看其能否比其他方法更“严谨缜密”。果然，人们认为有推广价值的唯一研究方法就在于人工智能研究方法的严谨性。这就好比，一个毕业生的作品是否被公开发表，决定了他能否受聘于一所名校。聘用后，他发表学术论文的数量，决定了能否获得联邦政府的研究资金继续研究人这项技术，而获得政府资助的研究项目是他获得终身教职的有利条件。

在过去的几十年中，神经网络研究屡次受到各个大学计算机系研究工作的重视，却又屡次被否定。致力人工智能研究的学者发现自己很可能要面临极具风险的研究生涯——曾有某个时期神经网络研究被认为是没有价值的，政府机构停止对该研究项目的资助。人们把这一段艰难的岁月称为“人工智能冬天”，神经网络研究人员也将面临艰难的再次职业选择——要么他们顶着艰苦的研究环境迎难而上，要么他们转变研究方向，选择那些国家财政有保障、专业领域获奖有把握的人工智能研究项目。

神经网络

在前几章中，我们探讨了人工智能软件之前发展中的局限性，软件通过把整个世界拆分成一系列“如果-则”子句集的逻辑函数表达，来尝试获取自动感知。而神经网络的研究采取了另一种方式。不同于逻辑驱动软件，神经网络以人和动物的神经系统为模板。虽然我们不能确切知道大脑如何工作，但我们知道大脑是由几十亿的神经元细胞所构成的。神经元是

神经系统的组成单位，同时在分布在脊髓和全身。

神经元细胞传递信号是一个电化学反应的过程。一个神经元细胞包含一个细胞体，负责接收信号，信号从其他神经元发出后要经过许多细小的树突（Dendrites）。当神经元细胞体接收到另一个神经元传过来的信号，信号会沿着一根长长的附属线路，类似蒲公英根茎的形状，它叫轴突（Axon）。神经轴突把一个信号传输给另一个。与其他细胞之间的连接点就叫作突触（Synaptic）。当神经元细胞释放电信号的时候，所释放到突触里的化学物质就叫作神经递质（Neurotransmitters）。

在大脑里，当神经元受到刺激时会“放电”，传递信号给别的细胞。视网膜神经元受到光线的刺激，它的电位会发生变化，释放出电能沿着轴突传输，把信号传递给突触，使其释放神经递质。其他的神经元接收到释放的神经递质，依次传递给其他的神经元，再依次往下传递。最后，经过一系列的激活，信号被传送到了位于大脑最高执行层的神经元，即前额皮层（Prefrontal Cortex）。就在这长长的一串神经元相互作用下，奇迹发生了，我们的意识告诉我们说，“看，这是我朋友！”或“小心，这是一只熊！”。

在这么一个没人能完全看见的过程里，我们大脑中的神经元学会了识别并对某一种特定刺激产生电位变化，至于其他不相干的刺激则没有反应。神经元之间的连接经过长时间的收到刺激或忽略刺激，会改变其连接性，即要么连接加强了，或者连接削弱了，我们把这个过程称为突触可塑性（Synaptic Plasticity）。一些大脑科学研究专家认为，如果大脑体验的是积极的过程，内啡肽（Endorphin）会加强活化神经元之间的连接。然而负

面的体验将会抑制或削弱这些连接。突触可塑性的概念在人工智能神经网络的研究中扮演着关键的角色，而且是整个深度学习的核心。

对于神经元细胞来说，最重要的问题是“我是否应该放电”。当神经元没有传递信号的时候，它处于休息的状态。当神经元周围的网络开始传输信号时，神经元可能在放电，也可能没有。神经元是否放电取决于所接受的信号累积强度能否达到阈值。如果信号强度没有达到阈值，细胞不会放电，一点儿也不会。

这一幅简单的神经网络工作流程图，就是生理上真实的神经元细胞工作的过程的简化版。然而，对于人工智能研究的目的，它提供了基础。现在让我们一起去看看人工神经元网络在计算机版本里所扮演的角色。

人工神经网络的研究没有被公认的确切开始日期，但20世纪40年代是该项研究的一个良好开端。1943年，两位电子工程师，沃伦·麦卡洛克（Warren McCulloch）和沃尔特·皮茨（Walter Pitts），发表了一篇关于为首个人造神经元细胞构建数学模型的论文，该模型被称为阈值逻辑单元（Threshold Logic Unit）。这个被他们描述为“大脑里的神经网络”简易模型，随即成为好几种人造神经网络的基础。

1949年，心理学家唐纳德·赫布（Donald Hebb）提出在神经元连接基础上加入神经元学习性的生理模型——赫布突触学习理论（Hebb's Synaptic Learning Rule）。赫布假定当神经元之间的激活连接被正向强化时（放大），大脑开启学习功能；同时，神经元连接被负向强化削弱的时候（缩小），大脑也在学习。赫布把大脑运行规律的新理论阐述在他的著作《行为的组织》（*The Organization of Behavior*）里。

提起赫布理论，人们通常指的是“共同激活的神经元成为联合”的说法。赫布这一对于神经元的简单推测随后被称为“赫布型学习（Hebbian learning）”。人们把赫布型学习理论应用到人工神经网络中，使计算机进行学习能力培养的科学也由此诞生。

今天在人工神经网络研究的领域包含了好几种不同的方法，但这些方法都有着一些得到普遍认同的核心概念。观点一：神经元彼此连接，组成一个去中心化的神经网络。观点二：当个体神经元细胞之间的兴奋加强或减弱时，大脑开启学习。观点三：只有达到一定的阈值，神经元才能放电或实现信息的传递。在人工神经网络的研究中，根据神经网络种类不同，设计需求也不同，这些公认的核心概念在应用时也有不同的呈现方式和细微的差别。

感知器

世界上首批人工神经网络出现在1957年，由当时纽约康奈尔大学伊萨卡学院心理学系的教授弗兰克·罗森布拉特（Frank Rosenblatt）提出。罗森布拉特把他制造的模型命名为“马克1号”。尽管制作的技术并不先进，使用了木质框架、成捆的电线、若干灯泡等，但是最终罗森布拉特制作出非常成功的感知器，“马克1号”就是现代深度学习神经网络的鼻祖。

感知器的出色之处在于其简洁性。按理推测，感知器可能是机器学习与模式识别最简单的实现方式。通过微调八个人工神经元细胞之间的连接

强弱，感知器能识别这些简单的模式。

罗森布拉特感知器的另一个特征在当时非常罕见，它使用的不是属于程序员难以改动的硬编码。相反地，这台机器运用了类似人类从经历中学习，能够灵活识别模式。这种方法在当时独树一帜，当时的人工智能技术里占绝对主流的可是模板匹配（template matching）与逻辑规则（logic rules）。

该感知器由六个机架构成，机架内装载着电子设备。每层机架相当于生物视觉感知系统中一个理论层，由上百个节点组成，最终达到了每个机架排列8个神经元。发展到今天，现代深度学习网络已经包含超过150层的人工神经元细胞，相互连接的节点有将近10亿个。这项伟大的创举实现了高速、强大的计算机和复杂的培训算法。

从生物学范畴来说，神经元细胞传输能力的强弱是由神经递质的集中程度决定的。弗兰克·罗森布拉特认为，化学神经递质的电模拟决定着连接各个人工神经元的输电线电阻的大小。低电阻（高导电性）对应着高神经递质水平，而高电阻（低导电性）对应着低神经递质水平。

感知器使用导线和导电晶体管来模拟人类大脑的结构。感知器的表层模拟生理上的视网膜，由400个光线传感器排成 20×20 的矩阵构成。（如果该感知器中的光线传感器的网格化布局让你联想到数码相机里的硅光电池传感器，请记住这种想法，这是一个极为重要的观念，我们稍后还会进行讨论。）



图9.1 弗兰克·罗森布拉特与一台连接着电视机外壳的感知器。图片展示的是感知器 20×20 的“眼睛”和连接A单元的导线

来源：Robert Hecht-Nielsen，“感知器”，加利福尼亚大学圣迭戈分校，0403号神经计算技术报告，2008年10月17日；照片由美国韦里迪安公司、通用动力公司提供

第一层神经元是具有传感功能的单元（sensory units），罗森布拉特将其称为S单元。

第二层神经元在生物视觉系统中模拟神经元树突。罗森布拉特把这部分称作联合单元（association units）或A单元。A单元由512个节点组成，节点与第一层神经元的光线传感器相连，每个节点看作一个突触。

第三个机架是感知器的输出层。这一层对应的是生物系统中的响应神经元，罗森布拉特将其称为R单元。它由8个灯泡组成（灯泡用以模拟神经

元细胞)，每个神经元都有特别的形状标记，如“三角形”或“方形”。只有接收到电信号超过阈值，神经元才会放电。R单元也遵循同样的原则，每一个R单元都有各自特定的阈值。当连接感知器光线传感器的A单元导线传输过来的达到一定的阈值，R单元就会放电，灯泡发光。

当感知器触发“三角形”的概念时，机器会渐渐学会“识别”该形状，并点亮贴有三角形标记的灯泡，公布它的答案。

究竟这个简陋的机械装置是如何识别模型的呢？首先，罗森布拉特用投影仪把一张图片投射到400个光线传感器上。图片是圆形、三角形或方形等某个简单的图形。图形其余部分是黑色背景。因此，图形部分比幻灯片的其余部分要明亮许多。光线传感器的光线据此做出反应。

当光线照射到传感器上，传感器会产生电压并把电流输往中间一层的机架，也就是A单元。光线传感器并非只是等待接收光线，倘若接收到从光线传感器传输来的电压的A单元达到一定数量，这些电压汇聚起来最终达到了R单元的阈值，引发放电。R单元放电，点亮对应着某个形状标记的灯泡。

无论是历史上或是现代的机器学习算法，都需要经过不断训练才能学会这套流程。在感知器的例子中，当发光的图形被投射到光线传感器上，有一个人类操作员坐在一旁指导机器回应。操作员先投影一张幻灯片，等待机器点亮一个或多个灯泡，然后为每个点亮的灯泡按下“对”或“错”的按钮。

整个训练过程需要煞费苦心地调整机器（或软件）人工神经元放电的阈值标准，以及连接神经元之间的节点强度，也就是专业术语所说的“突触强度（Synaptic Weights）”。如果机器回答不正确，人工神经元之间的节

点放电减弱。如果机器回答正确，人工神经元之间的连接要么自由放电，要么以某种神经网络的形式得以加强。

为了更好地训练感知器，每次操作员按下“错误”的按钮，感知器就会接收到一个很差的评级：A单元中通向错误灯泡的电线的电阻增大。几次训练过后，机器就不再轻易连上A单元中的部分导线，避免得出错误的答案。那么，下一次当A单元再次接收到从第一层光线传感器过来的电流，R单元的人工神经元达到阈值的可能性降低。由于R单元没有达到阈值，不能放电，在下次，系统就能识别出这是放电可能性较低的模式。如果机器得出正确的答案，则不需要做任何的调整，就像学生得到老师的肯定：“干得好，继续努力。”

一台全新的、未经过任何调整的感知器，得出的答案大部分一定都是错的。整个训练过程需要不断重复，循环展示同一系列的图片，直到机器回答完全正确。在后来几十年中，神经网络研究人员会围绕这个主题进行大量不同的试验，尝试包括在机器回答正确的情况下强化连接，改变电线的连接方式，增加神经元的数量或重新排列节点。

人类智慧取得成功的秘诀在于“练习，练习，再练习”。无论人工神经网络接受何种形式的训练，机器学习也有相似的秘诀——“重复，重复，再重复”。为了学会识别圆形与方形之间的区别，感知器需要不断重复，直到机器设定好应对每个错误答案的电阻。现代的机器学习仍使用同样的技术，但人工训练的过程被自动化、电子化过程所取代，这个过程以数据输入为基础，向软件中不断输入大量数据。



图9.2 ImageNet 2012年图片示例，相似的图片被排列在一起

来源：斯坦福大学Andrej Karpathy、Fei-Fei Li

相比之下，感知器的识别指令表是公认的非常有限的。与第四章中提到的移动机器人沙基类似，这款1959年的感知器也只能在非常有限的参数范围内运作。然而不同于“沙基”的是，该感知器的人工智能并非出自人类程序员之手。该感知器的机械系统是从人类“老师”的正误反馈中学习，摒弃了编程的方法，没有通过严谨的“自上而下”分析算法来分辨，比如说针对区别方形与圆形的算法。

当时，人们把感知器视为一项巨大的成功，向大众媒体宣传。罗森布拉特声称感知器能学习、掌握任何知识，掀起了轩然大波。《纽约客》杂志称赞感知器是一项重大的技术成就，1958年的《纽约时报》甚至把感

知器称为一场革命，刊登了以《新型海军设备在实践中学习》（*New Navy Device Learns by Doing*）为题的文章。

感知器为罗森布拉特在人工智能发展史上赢得了一席之地。在我所教授的机器学习导论课程中，第一份作业就是重现罗森布拉特机器的软件。罗森布拉特的成功及其所带来的大众媒体的狂热，在竞争激烈的思想流派中再次激发了计算机科学界的愤怒。

第一次人工智能冬天

如果人工智能这一全新领域有越来越多的研究人员遵循罗森布拉特的方法，我们也许能早几十年在自动感知领域取得成功。然而，取得良好的开端后不久，感知器便失去了光环。原因之一是训练神经网络所需要的计算能力和传感数据仍然不足。原因之二是人为问题——政治反对。

在罗森布拉特和感知器的故乡伊萨卡（Ithaca），冬季黑暗和漫长。然而，没有一个伊萨卡冬天能比得上罗森布拉特取得首次成功后遭遇的“冬季”那么严寒。具有讽刺意味的是，反对者之首是罗森布拉特的高中同学马文·明斯基（Marvin Minsky），被称为“人工智能之父”的麻省理工学院计算机学教授。正值媒体为感知器的自主学习能力大肆欢庆之际，明斯基却公开质疑罗森布拉特的观点——感知器能学会任何技能。明斯基坚决否认这种说法的真实性，认为“感知器没有科学价值”。

人工智能研究的舞台从此拉开了两大主流学派长时期的理论之争：符

号主义，主张由程序员编写代码搭建逻辑模型；神经网络，以机器学习为特点，通过数据驱动方法实现。作为人工智能研究资金的主要来源，军队一直密切观察哪一方在两大阵营斗争中占据上风。唯恐某一种人工智能范式垄断各界，那些为符号型人工智能呕心沥血的研究人员纷纷加入这场战争，坚信神经网络研究势必衰亡。

1969年，明斯基向罗森布拉特发起最后一击，携手麻省理工学院同事西蒙·派珀特（Seymour Papert）出版了《认知器演算法》（*Perceptrons*）一书。这本著作内容丰富，其中提及一项数学证明，旨在证实“感知器无法学会人工智能领域中任何一个简单的模式”：异或运算模式（XOR pattern）。就像音乐家使用C调音阶对乐器进行测试和热身，人工智能研究者通常用异或运算模式来测试机器学习算法。乐器出现任何的跑调情况都将如实反映在C调音阶上。如果连简单的异或运算模式都学不会，这就说明机器学习效率极低。

异或运算模式展现的是，所谓的逻辑互斥操作。如果两个字节不相同，操作员返回“1”；如果两个字节相同，则返回“0”。

明斯基和派珀特并没有真正地把异或运算模式应用到罗森布拉特的感知器上。他们采用数学建模的办法。他们用并不复杂的证据证明，不管经过多长时间的训练，罗森布拉特的感知器都无法学会模式识别。因为这个建模结果，明斯基和派珀特宣称，感知器无法学会识别任何模式，就连基础的模式也无法识别。

争论双方都有道理。确实，诚如明斯基指出，如果只凭借自己学习，感知器所能学会的非常有限。同时罗森布拉特也并没有宣称感知器能学会

任何技能。明斯基从字面上理解罗森布拉特的话，实际上罗森布拉特指的是一种理论化的、潜在、未来的机器学习能力，机器是由多层感知器层层叠加而成，而不指他所制造的“马克1号”。

具有讽刺意味的是，将近15年后，罗森布拉特的预言成真：计算能力的提高使神经网络的结构越来越精密，包含数十个电子层，层与层之间的连接节点高达数百万个。遗憾的是，在那个年代，罗森布拉特缺乏成熟的技术支持来回应明斯基和派珀特的挑战。多层感知器网络本可以证明罗森布拉特的正确，可在当时他的“感知器培训算法”只适用于单层感知器。

在这场激烈的理论之战发生的20世纪60年代末期，当时的计算机不仅体型庞大，而且计算速度缓慢。明斯基的理论获得了联邦科学基金会的认同，把明斯基的“自上而下”符号型人工智能范式视为最佳途径。结果，神经网络研究的资金被停止。

感知器项目戛然而止，所有曾经从事神经网络工作的研究员在这恶劣的学术环境下都不得不半途而废。此后不久，1971年夏，罗森布拉特于生日当天，在切萨皮克湾的一次船只事故中去世。他英年早逝，年仅43岁。而其高中同学兼竞争对手，明斯基日后成了符号型人工智能技术领域的发展先驱，研发出随后几十年内在人工智能领域独领风骚的技术。

讽刺的是，明斯基此后走上了罗森布拉特的道路，并在若干年后向媒体夸夸其谈人工智能技术的潜力。1970年的《生活》（*Life*）刊登了明斯基访谈内容，随后引起一阵轰动。明斯基发表了一次臭名昭著的声明：“在3~8年内，我们将研发出达到人类平均智力水平的机器。我指的是，机器人能阅读莎士比亚的作品，洗车、搞办公室政治、讲笑话、打架。到那时，机器能开

始以良好的速度进行自学。几个月后它的学习水平将达到天才级别，再过几个月，它的能力将难以估量。”尽管随着时间的流逝，明斯基的想法被证明完全错误，但在当时的几年内这种说法还是兴盛一时。

在麻省理工学院的机器人博物馆里，明斯基被视为英雄，具有远见卓识。然而由于他的固执阻挠，破坏了人们对早期神经网络的探索，把人工智能研究带入漫长黑暗的“死胡同”长达几十年。20世纪70年代是符号型人工智能发展的黄金时期。大量的研究人员通过编程方法投入仿真人类智商研究，这也凸显了人类对大脑工作原理的无知。

神经网络研究的复兴

符号型人工智能高速发展的同时，神经网络研究仍旧在艰难起步。到如今，“感知器”已经演变为神经网络中某一层神经元的通称。有趣的是，模拟硬件装置的命运与原始的感知器相似，已经被搁置一边。人工智能专家专注于在软件里打造神经网络。

1975年，哈佛博士生保罗·韦伯斯（Paul Werbos）发明了一种新型、改良的感知器。韦伯斯为推进人工神经网络的发展做出了两项重要的贡献。第一，他推出的神经网络不仅能输出“1”和“0”，还能输出分数数值。例如，0.5的计算值意味着神经元对该答案的“无法确定”。与罗森布拉特的感知器相比，那台原始的“马克1号”只能纯粹输出“1”和“0”，灯泡所能以“点亮”和“熄灭”的方式表示只能“是”或“否”，没有介

于两者之间的表达。

第二项贡献是韦伯斯提出了一项新的训练算法“误差反向传播算法（Error Backpropagation）”。这样一来，人工神经元能够以分数的形式来处理各种不确定性，误差反向传播算法用于训练多层神经网络的共同运作。罗森布拉特感知器的主要限制之一是输出的神经元层只能输出“是/否”两个答案，导致学习曲线过分陡峭，拖延进度。韦伯斯的新一代人工神经网络能使多层神经元联动，大幅度增加了其模式识别能力，也因此拓展了其潜在应用范围。这种新结构有利于提高神经网络处理复杂的模式分类问题。事实上，韦伯斯最先解决的难题就是此前一直被认为神经元无法破解的异或运算问题。

假设我们使用韦伯斯的方法来训练神经网络识别各种狗。首先，我们把一张 20×20 网络像素的狗的数字图片呈现给神经网络。传感层把信号发送给由400个人工神经元组成的首层神经网络，首层接收后通过加权连接把信号传给第二层，第二层再次通过加权连接把信号传给第三层，如此类推。

最后，当信号经过所有的中间层以后，抵达输出层。输出层由单一神经元细胞构成。该神经元细胞回答道：“这是一张狗的图片吗？”在能实现分数回应的前提下，网络能根据答案的确定程度，在0~1的范围内给出答案。例如，神经网络给出的答案是0.9，这意味着它比较确定图片的中的物体是一只狗，但不能百分百确定。

整个训练过程大致如此。如果网络提供的答案不正确，韦伯斯的算法会计算出哪些连接算出错误答案的概率最高。这部分连接将重新调整。相比罗森布拉特的感知器，通过改变导线的电阻来调整A单元（连接加权）。

而在以软件为基础的神经网络中，软件是通过改变其出错的连接的加权系数进行调节。通常，人工神经网络的连接权重是以百分比的形式呈现的。如何调节该百分比在任何一项训练中都是很关键的因素。

如果回答正确，韦伯斯的算法相应地算出哪些连接的正确率最高，然后通过增加权重来加强这部分连接。整个训练过程将不断重复，提供各种狗的图片以及其他的干扰项（除了狗以外其他物品的图片）给感知器判断。

尽管有了韦伯斯这套成熟的训练算法，神经网络依然面临着极大的限制。大的神经网络能够通过数学的方法保证正确识别所有呈现的图片，然而，当遇到新的、不熟悉的狗的图片时，该网络的识别能力将被严重削弱。神经网络能分别狗的图片和其他不相干的图片，比方说一座桥。但倘若遇上其他四只脚的动物图片，网络的表现将会大打折扣，陷入混乱。这情形就像同学在考试中，遇到多项选择题时，胡乱组合答案。

然而，充满希望的春天如约而至。先进的数字摄像技术加上韦伯斯误差反向传播算法，在神经网络的研究里点燃了新的亮光，很快结束了20世纪60年代到70年代的人工智能冬天。如果你仔细研读80年代末至90年代的研究文章，你能感受在这期间，研究领域呈现一片欣欣向荣的景象。研究人员尝试应用神经网络对太阳底下的一切进行分类：图像、文本、声音。神经网络研究——现在这种行为被重新命名为“联结主义（Connectionism）”——应用领域十分广泛，从对信用卡申请人的评级到药物分析无处不在。

可是，20世纪90年代的神经网络研究复兴并没有持续很久。尽管无数的学术文章表明神经网络在人为结构设置的环境中能够取得许多成功，

但在实际应用中它却表现不佳。问题恰恰在于机器学习得太好了。神经网络的过度指定（overspecialize）、过拟合（overfit），仿佛一个小孩子能够准确无误地记住数学问题的答案用于应答，而没有理解答案背后的公示。

过拟合问题在于，神经网络能够学会识别训练过的图片内容，却缺乏“举一反三”、对知识内容扩展的能力。僵化的不变量问题依然没有得到解决。如果人们训练神经网络识别100张猫或狗的照片，它可以成功地对这100张图片正确分类。然而遇到训练内容以外的图片，神经网络将很可悲地遭遇失败。这使得联结主义学界再花上个十年研究，如何在机器出现过拟合之前停止训练。这种方法叫作“提前停止法（Early Stopping）”，只可惜这个方法来得有点晚。

这一次，对神经网络造成打击的不是人工智能界的逻辑主义阵营，而是来自同一阵营的竞争对手阵营——新的机器学习方法。那个时候，这种机器学习方法，特别是支持向量机（Support Vector Machines）开始盛行。他们相比神经网络的表现进步并不大，却十分诱人。对于视觉识别任务，比如那些对无人驾驶来说十分重要的识别需求，机器表现提高了1%——若是在以秒为单位的、激烈竞争的奥运比赛场上，任何细微的差别都可能扭转赛果，那里是“胜者为王”的赛场。

支持向量机的大部分进步之处，回想起来，也只是能在路径清晰的发展中避免过拟合的问题。这门技术被称为“常规化（Regularization）”。当同样的过拟合缓解办法应用到神经网络上时，虽然神经网络也能有所提高，可惜这一切太晚了。神经网络研究在20世纪90年代中期再度失宠，对不变量表示（invariant representation）的问题仍旧没有得到解决。

神经认知机

在神经网络的发展历史上还有另一段畸形扭曲、颇具讽刺意味的传奇故事。它发生在20世纪90年代短暂而风光一时的人工智能研究复兴之前，80年代的一款神经网络模型实际上解决了不变量表示的老大难问题。在详细讲述该神经网络的学术文章刊出以后，引起了许多神经网络研究“顽固对头”的关注。

1980年，日本科学家邦彦福岛（Kunihiko Fukushima）提出了一种新型的神经网络模型：神经认知机（Neocognitron）。这种新型的神经网络机器使用了许多之前深度学习网络并未应用的新技术，这些技术直到今天在图片识别方面依然出色。然而，在这样一个已经被人们研究多年的领域，福岛的天才发明却缺乏关键的使能技术来引起世界关注。

福岛的目标是创造出多层神经网络，能够稳定地进行视觉识别。更准确来说，他的目标是创造一款机器学习软件，用于辨识手写的邮件，并自动分类。

福岛的网络软件有三大主要且“及时雨般”的重大进步。首先，当图片被“喂食”到网络时，并非一步到位，图片是被分成若干组像素群，而不是把整张图片喂食到输入层。这样，第二层的每个神经元并非与输入层的所有图片接收器连接，实际上每个神经元细胞只是与若干像素连接在一起。这些像素之间产生轻微的重叠，这种“滑动窗口（sliding window）”的

想法被应用在各个神经元细胞层里。

第二大进步是神经元克隆。福岛没有使用大量的神经元，而是每一个神经元有着各自的连接。神经认知机每一层的神经元都是克隆的产物，突触连接的力度相同。在这样的前提下，尽管每一层可能有成千上万的突触连接，只需要几个参数就能设置连接的力度。这种构想来源于对数字图像的观察，识别图片左上角的字节的神经元细胞，与识别左下角的字节时的神经元细胞是同一个。换言之，图片的不变量表示意味着网络当中对应的不变量。这样，这款神经认知机的训练进度就将快得多，也平稳得多。

最后一项改进是由两种类型的神经元组成的网络人工神经元：S细胞用于特征提取，C细胞用于包容这些特征的变化情况。当不同的神经元层交替运用时，就会应用到上述神经元。

福岛的神经认知器采用了韦伯斯的误差反向传播算法。当某种刺激图形（stimulus pattern）呈现给传感层时，该图形就会被分成几个补丁程序。一个常见的补丁程序包含9个传感器，排成 3×3 矩阵与下一层的神经元相连。

在该神经网络中，人工神经元S细胞与C细胞排在重叠的补丁中。重叠的部分是为了确保图形的每一处特征都能被神经元感知到，倘若某处特征没有被准确感知，那么马上启用旁边的神经元。S细胞与C细胞的交替在一起工作，渐渐把碎片化的信息整合成完整的视觉图片信息。每一层细胞都按高低阶层有序排列，下一层能够看到前一层的神经元细胞。那么，当每一层神经元细胞逐渐把碎片整合成完整的图片信息，最里层的神经元也能间接感知到整体视觉效果。

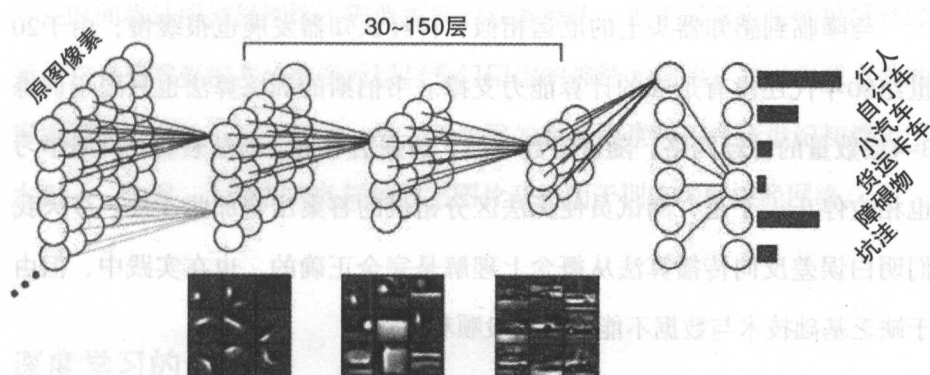


图9.3 福岛的深度神经网络的结构示意图，圆球代表个体神经元。同一层内的所有神经元都是经过克隆得来的，完全相同。图中显示的是个体神经元在训练中接触到汽车图片时产生的视觉特点。最左边的特征代表基本结构；中部特征代表车辆的局部，如车门和轮胎；最右边的特征代表车辆的整体情况

来源：国际计算机学会举办的2009年第26届深度学习国际会议，由Honglak Lee、Roger Grosse、Rajesh Ranganath和Andrew Y. Ng合著的《卷积的深度信念网络关于图像分层的可扩展、无监督学习》（*Convolutional Deep Belief Networks for Scalable Unsupervised Learning of Hierarchical Representations*），609~616页

在神经认知器之前，包括罗森布拉特的感知器在内，历史上所有的神经网络，都把完整的视觉场景一次性输入到第一层神经元网络，事实证明数据过于庞大，无法得到妥当的处理。

神经认知器是一项奇迹。它独特的结构足够稳定，它能识别在手写信件中每一个字母，轻松识别字母位置的变换。神经认知器的模式是现代深度学习网络的“老祖宗”。它在同时代的图片识别模式中脱颖而出。不久福岛的构想被扬·勒丘恩（Yann LeCun）、约书·本吉奥（Yoshua Bengio）和其他研究人员所借鉴，创造出我们今天的卷积神经网络（convolutional neural networks, CNNs）。“卷积”这个术语指的是克隆某项数学函数，并不断将其应用到重叠的常规矩阵中。

与降临到感知器头上的厄运相似，神经认知器发展也很缓慢，由于20世纪80年代还没有足够的计算能力支撑。韦伯斯的训练算法也只限于训练3~4层数量的神经网络。随着层数增加，加强信号也会渐渐衰弱，网络学习也相应停止，于是，测试员便无法区分错误的答案出自那些连接。今天我们明白误差反向传播算法从概念上理解是完全正确的，也在实践中，但由于缺乏基础技术与数据不能如设想般顺利运作。

1990-2000年期间，一些学者尝试采用“浅层化”的仅有两层神经元网络来弥补技术与数据的不足。对于神经网络的数学理论层面来说并不起什么作用，但万能逼近定理（Universal Approximation Theorem）受到了关注。万能逼近定理认为，至少从理论上来说，神经网络实际上只需要一层“隐形的”神经元，用于任何可量化的函数中接近所需达到的准确度。换言之，根据定理，从数学范畴来看，浅层化网络能够不受任何理论限制，必将取得成功。

另一方面，万能逼近定理惠及神经网络发展。毕竟，这是一个好消息，神经网络的整体核心概念是行得通的（至少从数学上看），也具备投资价值。另一方面，数学证明再次误导了研究人员，他们向错误的方向疾驰而去。他们尝试证明不增加层数，只改进训练算法的办法，来提高隐形的神经网络层的性能。

由于20世纪80年代到90年代的技术限制，许多专家非常遗憾地认为万能逼近定理意味着尝试建造超过两层的神经网络没有意义。这就是一个人投入宝贵的时间和资源用于建造多层神经元网络。乐观来说，研究有风险；悲观来看，纯粹是浪费时间。

时间将证明神经网络不需要更好的训练算法，也不需要大量的单层神经元。它只是需要很多的神经元层与专门用途的神经元细胞，并以串联重叠的形式排列。多亏了社交媒体，另一种急需的重要因素终于在新世纪初浮出了水面——数据，大量各式各样的数字图片开始用于训练深度神经网络。

深度学习的诞生

如同新生儿，机器通过接触大量的信息进行学习，即数据密集型办法。在训练算法中，所需的数据数量通常与问题的难度成正比。相比之下，教会一台机器辨别三角形与正方形，比教一台机器成功辨别男性与女性要简单得多。罗森布拉特的感知器只需要调整512处节点的权重，所以用于培训的图片数量相对较少。大型的神经网络（无论深层与否）都包含上百万的连接，因此需要上百万张训练图片。

如果把算法比作引擎，数据就是所添加的汽油。引擎如果缺少汽油，是无法运作的，同理，没有数据支撑的机器学习算法就是个不中用的摆设。

20世纪，大部分机器视觉的研究人员都承认数据对于训练的重要性。在计算机发展历史上看，绝大多数情况下，数字图像是很难获得。于是，图像识别算法的发展使用了“数据稀疏法（data-sparse approaches）”。由于数据稀有，机器的学习训练就只能演变成类似生物界的物竞天择——必须适应食物稀少的环境一样，只能集中精力发展运作高效、使用对数据数量要求不高的算法。

数十年来，高效的机器学习算法的需求仿佛把研究人员带进了《爱丽丝漫游仙境》里的兔子洞，进入另一个满是询问和微调的计算机算法世界，他们努力地有限的数据集争取1%的性能提升空间。同样的想法也能推论出人设计的正确算法比受生物界启发的办法更有价值——人类的认知能力通常伴有充足的数据与快速、大量的平行计算等特点。

在生物学里，受重视的不仅仅是计算效率，还有适应性和健壮性。一个生物的命运取决于其神经算法能否快速适应新的环境。倘若我们的大脑中果真运行着一系列的学习算法，至少从传统的计算机科学场景来推测这套算法很可能结构简单、效率不高。

从数据贫乏的“智能”算法到数据充裕的“简单”算法，这种转变发生在2010年之后。以下的科技发展共同推动了该转变的发生：计算机价格下降、运算速度提高；手机开始兼具数码相机功能；互联网给予人们足够的空间存放所拍摄的数码照片。要知道，Facebook每一分钟就将出现208,300张新图片。再见了，数据匮乏的年代。我们迎来了谷歌图片搜索的时代。

当计算机图片搜索盛行后不久，各种信数据信息的“大坝”面临着“决堤”的风险。2003年，一位名叫李飞飞的加州理工学院学生创造出CALTECH 101，一款图像存储器，能把9146张图像拆分成101种不同类型。李的目标是建立图像合集，图片包含各种人类日常生活的场景，用作机器视觉训练算法的素材。

到了2006年，李飞飞的数据库发展成为CALTECH 256，能存储30,607张图片，分成256种不同类别。2009年，在伊利诺伊大学香槟分校和普林斯顿

大学待了一段时间后，李飞飞加入了斯坦福大学计算机学院。尽管她听到热心同事的好心建议“去研发一些更有用的东西”，她还是决意继续搭建更大的数据分享平台ImageNet。最终，ImageNet发展成为全球首家图像大型数据库，存储有过百万张不同标记的图片。并且斯坦福的ImageNet的资源一直持续增长。截至本书写作之时，ImageNet的图片总数超过1400万张，2万多种分类。

ImageNet上的图片是原始粗糙、杂乱无章的。物体毫无规律，如果是发生在现实生活中，需要获取上下文信息才能理解。比方说，一张啤酒瓶的图片，可不是那种在广告上呈现的光鲜亮丽的啤酒瓶，而是一堆被乱扔在人行道上的啤酒瓶子的模糊影像。

ImageNet包含着各式各样不同种类的图片。这其中必然包含有诸如猫和狗这样的常见图片，还有蜥蜴、蜗牛、蛇、雪橇和袜子等图片。另外，还有一些古怪有趣却略显业余的抓拍，如道路、烤箱和番茄等。

李飞飞的目标并非囤积一堆杂乱无章的数字图片合集。这些照片都有各自的分类。当互联网上全是各式各样的图片时，需要有人认真查看图片内容，恰当分类；此后才能用于神经网络的视觉训练。人类必须亲自查看上百万图片，创建相应的分类。必须有人愿意耗费时间承担这繁重的工作，翻阅一沓狗的照片，弄清楚照片X上是约克郡犬，而照片Y是约克郡犬的近亲——斯塔福德郡斗牛犬。

这时恰好出现了一股文化潮流有望解决这个难题：众包。对ImageNet上百万张图片进行分类，这个极其艰巨的任务最终由上百名亚马逊土耳其机器人（Amazon Mechanical Turk）完成。它们为每一张图片标记，并为每张

图片的工作量付费。据李飞飞回忆，ImageNet曾一度是亚马逊平台最大的自由职业雇主，并在全球范围内全天候招聘员工，数量多达几千名！

绘图处理器（GPUs）

我们已经对高速计算机、大数据、数字相机和多层人工神经网络等主题进行讨论。还有一项更为重要的技术即将映入眼帘：高速独立显卡。它之所以重要的原因在于神经网络对于计算能力的要求极高。

早先，优质的独立显卡资源掌握在游戏界手里。游戏行业有时会遭遇某些保守的成见——“游戏会使青年人沉溺其中，浪费时光，毫无创造性”，然而事实上游戏行业却是创新的主要来源，尤其在图片处理方面。

游戏能测试出计算能力的高低。在游戏中，计算机需要以高清像素和快速帧率再现3D图片场景。计算机必须实时响应用户的输入，抵消几个游戏玩家之间的网络延迟时间，同时通过四射的水花和椅子的摇晃形成精准的肢体刺激。而且游戏应用对于计算机环境的要求更高，比普通应用程序的要求（如复制电子表格里的数字或数据搜索）更复杂难度更高。

连摩尔定律的速度也无法满足游戏行业对于低成本、高性能的计算机要求。于是，游戏行业只能采用了另一种方式来突破瓶颈：平行启动多个处理器。这致使游戏的硬件制造商没有研发更快速的处理器，而是研发专门的独立显卡，包含上千个平行处理器。

独立显卡有大量被称为GPUs的平行处理器（为了与传统的中央处理器

CPU区别开来)。GPU_s被看作是连接图片设计与游戏的特别应用。CPU的性能提升曲线可谓增长迅猛。而相比之下，GPU_s的性能提升过程发展更加神速，该趋势引起了极大的关注。2006年，全球最大的独立显卡生产商之一的英伟达(Nvidia)引入了GeForce 8系列，这款GPU经过特别的设计，用途不局限于图片处理。

为了打开新产品市场，Nvidia创造了新名词“通用计算图形处理器(general purpose GPU, GPGPU)”，宣传这是台式计算机平行发展的全新领域。在Nvidia引入GeForce 8系列之前，只有专业的图片艺术家和科研精英才会用到通用平行计算。现在GPU拓展到那些需要高强度平行计算能力的应用领域，比如股市交易、工程分析，当然少不了神经网络。

神经网络正朝着平行化的趋势发展。神经网络包含许多的神经元，每个神经元处理各自接收到的输入内容，决定是否放电进行输出。这种情况非常适用于多层卷积网络，因为其包含的阵列组织能够与图像应用结合起来，这也恰好是设计GPU_s的最初目的。

在大多数的计算机应用中，速度是非常重要的，对于神经网络而言，速度更是重中之重。虽然神经网络进行图片分类时，其在GPU上计算出答案的速度仅比在CPU的速度要快上一两秒，但当神经网络在训练时，GPU_s的优势就十分明显了，因为训练需要进行数百万次的反向传播迭代。使用GPU_s运作卷积神经网络的研究人员，相比其他使用台式电脑的同事，都能够觉察到一个明显的改善。原则上来说，对于大多数应用程序来说，训练过程的速度并不是那么重要。毕竟，训练往往只需要进行一次就够了；而在实践中，现实世界没有人愿意研发、编程一套训练时间长达一个月之

久的算法。如果人们面对的是一个只需三天时间便能完成训练的网络，那么情形将大不一样。

现代深度学习

2012年，现代深度学习在一场图片识别大赛上表现亮眼。大赛的目的是妥善利用ImageNet上的分类图片的宝贵价值。2010年，李飞飞与同事开始创办ImageNet图片识别竞赛，一个每年一度的大规模视觉识别挑战赛，任何人都可以参加。

比赛规则如下：参赛者向大赛管理方监管的服务器提交图片识别软件。根据要求，软件将会处理十万张新图片。由于图片随机排列，呈现的物品分散，软件将需通过编程和算法，对每张图片上五种识别出的物品进行命名。

ImageNet竞赛分为三项：图像分类、分类定位、检测。分类项目旨在测试计算机算法能否将图片正确分类并标上恰当的标记；分类定位项目用于评估对已标记图片以及图片上的物体位置的复制能力。最后的检测项目借鉴了其他项目的元素，但评价标准更为严格，图片数量众多，每张图片上呈现好几样微小的物体。随着时间的发展，比赛会新增了一些项目，如视频流的识别，保证比赛跟上技术的发展。

2010年比赛的获胜者是来自日本电气股份有限公司（NEC）和伊利诺伊大学香槟分校。在10万张测试图片中，获胜的神经网络的出错率是28%。

两支亚军团队的错误率分别是33.6%和44.6%。为了更好地分析这个比赛结果，我们来对比一位未经过特别训练的普通人，他对图片进行分类的能力相当突出，错误率只有5%。

2011年第二届ImageNet竞赛成绩显示，尽管参赛团队几乎都没有使用神经网络，但这一年中机器视觉算法还是取得了较大进步。来自XRCE技术公司的团队错误率只有25%，比上一届冠军降低了2.4%的出错率，两支亚军队伍的错误率分别是31%和36%。

当第三届ImageNet竞赛在2012年9月30日深夜落下帷幕时，机器视觉领域发生了永恒的改变。这届比赛既没有狂热的媒体，也没有华而不实的颁奖礼。但是如果监管比赛的机架式服务器有反应的话，它们将会彼此拥抱，热烈庆祝。

东京大学团队研发的SuperVision神经网络，正确率高达85%，创下了世界软件识别比赛中骄人的成绩。错误率降低15%~25%听起来并不太多，可是对于计算机-视觉研究界而言，他们过去苦苦努力为了提高不到1%的正确率，这成绩就像是看到一个人人生第一次达到专业田径选手的水平，四分钟内跑完一英里。

SuperVision的发明者是亚历克斯·克里泽夫斯基（Alex Krizhevsky）、伊利亚·斯图斯盖尔（Ilya Sutskever）两名同学和他们的老师戈弗雷·辛顿（Geoffrey Hinton）。Super Vision属于卷积神经网络。卷积神经网络的许多特征是建立在三十多年前福岛博士发明的神经认知器的技术基础上的。其余的改良部分是来自于由纽约大学扬·勒丘恩、斯坦福大学的吴恩达（Andrew Ng）和蒙特利尔大学的约书·本吉奥组成的研究小组的工作成果。

SuperVision是一个大型的神经网络，由工人神经元组成矩阵，矩阵组成机架，机架再组成一个30层神经网络。东京团队做了一个大胆的举动，他们开放码源，让任何人都能使用和修改。这个举动在计算机视觉界内轰动一时。

深度学习人工神经网络已成为图片识别软件新的黄金准则。在SuperVision取得2012届图片识别比赛的辉煌胜利之前，几乎没有人使用卷积神经网络。但在2012年后，不使用卷积神经网络的研究团队甚至不敢再参加竞赛了。

SuperVision获胜后的下一年，冠军的错误率降低至11.2%，亚军紧跟其后，数据分别是12%和13%，所有人都使用的是指定的深度学习卷积神经网络。2014年，谷歌团队创下了6.66%的错误率新低。来自牛津大学的团队使用了更庞大的卷积神经网络，但错误率是7.1%。2015年，微软北京研究实验室团队（由首席研究员孙剑带领）使用了152层深度神经网络，赢得了三个比赛项目的冠军。值得一提的是，微软团队这一年的错误率仅有3.57%，历史上首次低于人类的错误率5%。

在取得这些胜利之后，机器视觉其他的研究方法仿佛都要被淘汰。算法不再局限于物体识别，已经开始从计算机视觉领域扩展至所有人工智能领域。人工智能发生了永恒的改变，无人驾驶汽车发展的最后一道障碍——软件的人工感知能力——终于被扫除了。

这场巨大的胜利后不久，各方开始尝试合作。Nvidia 推出了深度学习卡，是克里泽夫斯基的SuperVision网络在低功耗硬件上的衍生产品。Nvidia 的商业应用程序目标何在？毫无疑问，是无人驾驶汽车。该系统被名副其实地命名为DRIVE PX，它能同步处理超过12个视频频道的实时内容。一年

后，速度更快、造价更低、质量更优的计算机卡推出，自动深度学习的技术研发竞赛拉开了大幕。

神经网络内部

有几种不同类型的网络用于图像识别，每个网络都有自己的自测工具，并通过应用独特的改进算法提升训练的精度。深度学习是快速发展的领域，几乎每周都会推出新的结构和算法。然而，一个共同的特点是深度学习网络使用串联式多层人工神经元，从经过软件识别与标上标记的数字图片中捕捉图片特征。先进的深度学习网络拥有超过100层的人工神经网络（对比罗森布拉特的感知器只有单层神经网络，共八个神经元细胞）。

有人认为深度学习网络识别物体的方式与人类相同，首先认识某个细小的特征，然后把该特征抽象出来，应用至更广义、更抽象的概念中去。虽然直到现在，生物系统识别物体的过程仍然是个谜，但人类的眼睛看到一双尖尖的耳朵，八根胡须和毛茸茸的尾巴时，总能快速地把以上视觉信息分类：

“啊，是只猫！”是的，构建神经网络时，通过进行快速的个体特征分析确实是物品分类的一种办法，人工感知也由此更接近人类的感知水平。

让我们以SuperVision这款神经网络为例，看看这些多层神经元解析机（Multilayer Analytical Engines）是如何工作的。使用GPU加快训练过程是克里泽夫斯基及其团队作出的一项非常实用的改进，使得训练周期从以“周”为单位减少到“天”为单位。鉴于SuperVision是大型的神经网络，

神经元解析机数量多且密集，系统参数高达6000万，含有神经元数量65万个，为大幅缩短训练时间提供了极大优势。

在神经网络内部，SuperVision（研究人员为其取了小名“AlexNet”）采用了更为简单的阈值形式。经过简化的传输功能有助于神经元提升速度，同时使韦伯斯的训练算法深入多层神经网络调节其中的节点，有利于解决长期以来困扰多层神经网络的问题。为了解决过拟合的问题，SuperVision团队采用了一种名为Dropout的技术。Dropout技术包括在训练过程中清除部分连接不畅的神经元——没有任何单独的神经元能完成所有工作。Dropout能够让所有的神经元都各施其职参与到计算过程中。

创建深度学习网络的第一步是，提供原始视觉数据。每张数码图片包含一个矩阵，矩阵数值以红绿蓝光量化在每一个像素里。深度学习网络的输入层内有三个相似但互补的输入矩阵，数值能矩阵之间相互流通。

由于深度学习网络的种类不同，其神经元层的排列也各异。在标准的卷积神经网络中，第二层神经元与第一层具有数据功能的矩阵相连。例如，第一层含有 3×3 排列的矩阵，第二层的神经元负责计算出这些像素的加权总和——这与福岛的神经认知器极为相似，排列的神经元能够帮助神经网络完成性能调校——如果总数大于一定阈值，神经元将会“放电”或向下一层神经元发出特定信号。如果数值过低，神经元则安静等待，处于休息状态。当信号扩散后，或者说在神经元之间传送时，神经链中的下一个神经细胞将计算从 3×3 的窗口（window）里传过来的信号总数。这个过程在神经网络中一直持续进行着。

大多数深度学习结构中也包括最大池法（max-pooling）的单元，与神

经认知器的C细胞相似，从神经元池中抽取最大数值，忽略其他信号。结果证明最大池法能加强网络的稳定性。在大多数的深度学习网络的最后几层中通常包含两层及以上的老式“完全连接”的双层感知器。最后卷积神经网络学会并发现最佳图像特征，这时感知器开始运作。

当信号抵达最后一层的神经末端，输出神经元会统计内层神经元细胞的“投票结果”。例如，罗森布拉特的感知器就会以点亮灯的形式来表示最后的输出结果。如果一款现代深度学习网络的软件接受识别猫狗图片的训练，它将根据机器对物体的识别肯定程度，在数字0~1范围内表达输出的数值。如果在判断狗的时候，对一张图片的输出数值是1，那么这就表示神经网络百分百肯定图片上是一只狗。同样，假使判断猫的图片的数值是0.5，就表示网络不太确定图片上是一只猫。

显然，这种对深度网络学习分析法的解释极其简单。随着该领域研究的发展，大多数网络都对这个流程进行了改良。改良做法包括：把功能独特的神经元层插入中间层，进行实时数据分析，还原可能被扩散的信号。其他的改良之处，还包括2015年微软团队设计的训练方法“残差学习（Residual Learning）”——使训练覆盖到更多神经网络层。

深度网络学习的优点之一是如果设置妥当，网络能在对重复物品识别的基础上，自主形成对新数据的识别能力。具有讽刺意味的是，机器本身，而并非人类程序员，自主培养出对新出现的陌生物体的识别能力。深度学习网络是程序员口中的典型的“黑盒结构”例子，意思是当软件进行输出时，几乎不可能对软件程序的步骤进行任何逆向操作。

假使一辆无人驾驶汽车把眼前的一群过路行人，误认为是玻璃钢结构

的摩天大楼墙体反射的倒影，那么即使该程序的设计者也不会知道为什么视觉识别软件迟迟不能检测出错误原因。现代多层深度学习网络的神经层可以达到几十层之多，层内包含上百万甚至几十亿的连接点。就好比一个人不知道自己的想法有什么确切的逻辑依据，只是随口凭感觉来回答。深度学习网络也是同样的，它们是凭借上百万的人工神经元的“感觉良好”来得出结论的。

新型边缘检测

分析深度学习网络的识别行为的一个方法，是把人工神经元逐个分开，用某种特定模式来测试它们的反应。进行这项实验的研究专家发现，在深度学习网络识别视觉信息的过程中，人工神经元细胞在网络中的位置越深（越靠近输出层），识别抽象模式的能力越强。相反地，神经元细胞越靠近表层网络，识别出的物体越简单。

1959年，有一次有趣的生物学实验中，生物学家大卫·休伯尔（David Hubel）与托斯坦·威泽尔（Torsten Wiesel）记录了经过轻度麻醉的猫（也许还有其他动物）的视网膜活跃性——以尺寸不同的斑点和黑白条纹刺激猫的视网膜。当他们检查神经元对视觉刺激的反应时，他们发现这些视觉皮层细胞仅仅能识别条纹，不能识别点状物。于是，他们两人把这些细胞称为“条状物检测器（bar detectors）”或“边缘检测器（edge detectors）”。事实证明，深度学习网络的外层神经元对于线条和边缘反应更加强烈。

在深度学习网络神经元中，首先会将复杂的图像分割成单一的线条和边角，目的是要算出图片里究竟有些什么。在过去几十年中，例如“沙基”这样的早期机器人，是通过人工预先编程视觉信息分割成线条和边角，以便更好地感应物体。而到了今天，深度学习网络则采用自主化的流程处理方式。并且在深度学习网络中，越进入深处，人工神经元所回应的模式就越复杂。如果该网络训练时使用的是汽车图片的数据库，那么它的第三层神经网络很可能对于类似车轮的圆形物体作出判断。第四层神经元也许能识别汽车的某部分或某个种类，比如说卡车或车前挡板。再往网络深处走，神经元就能识别更抽象的概念，例如从驾驶员座位上看到的，与车外某个物体匹配的3D模型，等等。

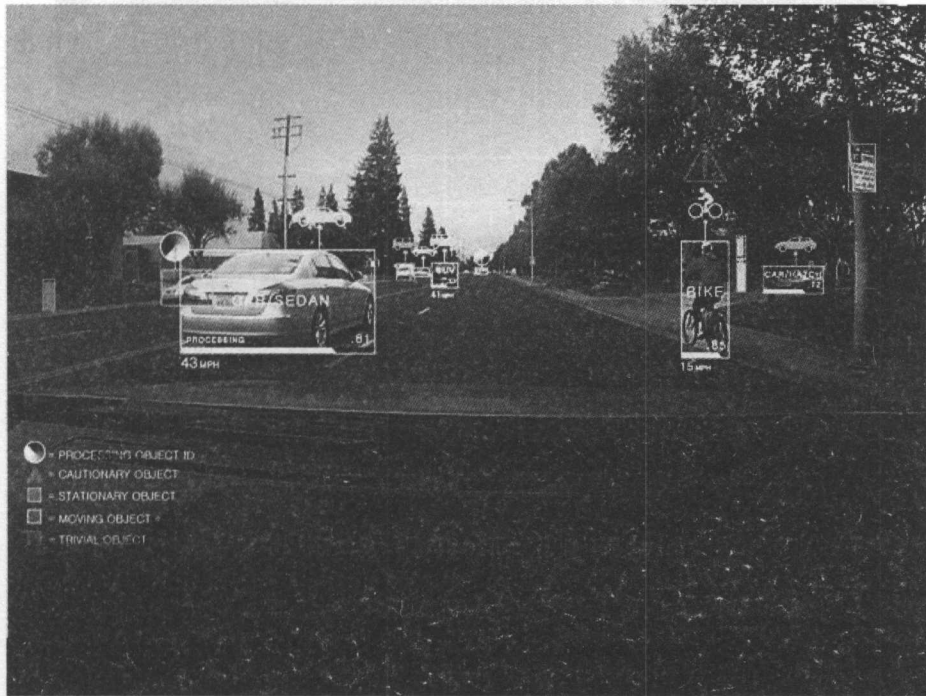


图9.4 深度学习应用在驾驶途中的实时物体识别
来源：Nvidia公司

当研究人员渐渐熟悉这种学习规律后，就开始将它运用到实际设计中。到后来，无论卷积神经网络是为了何种应用而接受训练，网络首层几乎都会包含边缘检测器，因为这种边缘检测技术在图片理解领域是普遍适用的。

在网络深处，内层神经网络开始朝“专业化”发展：受训后能识别汽车的神经网络能马上反应出车相关的特征。可能生成专门识别车身形状的神经元，以及其他具有识别特定部位的神经网络。在训练时学会识别狗的神经元能对“皮毛”“耳朵形状”等特征作出响应。

我们把引起神经网络响应的模式称为“视觉特征（Visual Features）”。对于初级水平，图片特征简单：线条和片状物。当水平渐渐提升，视觉特征变得更加复杂和抽象，甚至连“特征”这个词都不能很好地概括它的含义。位于高层网络的个体神经元能够仅通过部分视觉特征的组合，就作出判断。例如，某个神经元能对前格栅中识别出这是“雪佛兰汽车”。

哲学家们绞尽脑汁来描述个体独有的感官体验，却无法量化或直接表达比人与人之间的个体感受差异。专业术语“感受质（Qualia）”指的正是个体直接的感受体验。我们都明白自己的感受，但是无法确切比较自己与他人的感受异同。例如，每个人看到的傍晚天空的红颜色都一模一样吗？巧克力在你口中感受到的味道与别人所感受到的相同吗？

也许未来某天，深度学习网络能够拥有自己的感受质。多年前，我便亲身感受到深度学习网络发现自我认知存在时那种诡异，当时，我和学生正在准备一场深度学习网络的现场演示。我们每天都会用几小时，把能在实验室找到的各式物品随机在与神经网络联网的摄像头前展示。神经网络照常接受训练，并把大部分图片根据内容进行正确的分类。然而，奇怪的

事情发生了。我们意识到网络仿佛一个月的婴儿正在识别我们的脸部。每一次当我的头靠近摄像机，网络产生反应。当学生们效仿我的做法，网络仍旧产生回应。

我们发现在神经网络第七层的某处有一个神经元，每次人脸出现在图片上时，该神经元都能马上反应。在长期训练过程中，这个人工神经元培养出专门检测人脸的功能。我们并没有特别训练网络进行人脸识别，可为什么它能够识别呢？

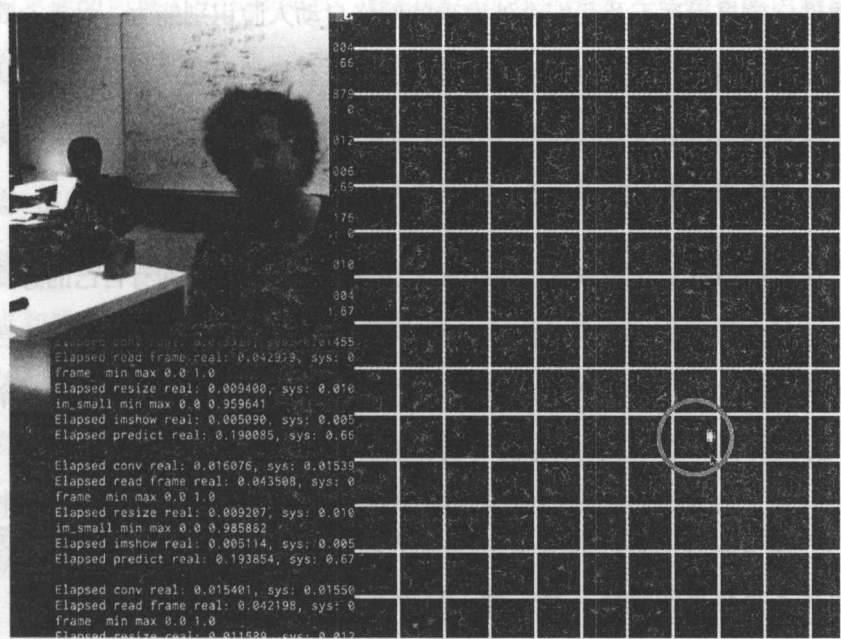


图9.5 这是一张意外拍摄的图片，卷积神经网络正在回应左图视频部分的两张人脸。右图上用圆圈标志的区域里有两个模糊的白点，正是对应左图框里命令中的两张人脸

来源： Jason Yosinki，康奈尔大学

我们疑惑了很长时间，后来我们认为人脸的出现对于神经网络识别物

体时是一个带有有用信息的字节。我们人类也有把某些物品（比如手机）靠近脸部的习惯。神经网络在仔细观察我们一阵子（每天几小时的训练）之后，神经网络作出决定要检测人脸，因为它（也许）认为这样才能更好识别我们训练展示的物品。

这是一个惊悚的时刻。

网络竟然被培养出了自主学习的决策能力，这个发现具有重要的意义。原因如下：第一，我知道许多同学花好几年时间在大学时期研发人脸检测软件，效果远远不及我们意外发现的网络自动人脸识别能力。另外，我们意外发现的网络自动人脸识别能力带出了一个有趣的问题：在我们固有认知的范围以外，网络还想要检测什么呢？很可能它想要识别某些重要的视觉模型，那些模型无法用言语描绘，甚至连人类的大脑也想象不到的。

这段经历不禁让我深思，是否某天深度学习网络就能拥有自己的感受质呢？想象以每秒1000张的速度，投放图片来训练网络。想象一次使用十台摄像头。也许还有其他形式的传感数据，比如超出人类听力范围的声学频率数据。我很好奇，当这一类的训练开展几周后，机器究竟会发生了什么？

从理论上说，软件能培养出人工神经元回应感受质的能力，而感受质是超乎人类想象的物体，是一种“只可意会不可言传”的状态。显然，人们轻看了机器智能的能力，常断言计算机永远无法体会生活中的美好，比如欣赏落日余晖或品鉴红酒的芳醇，我想，“这话不错，可计算机自己独特的传感体验，也可能是我们从未知晓的”。

深度学习网络为未来人工智能研究创造了沃土。与人类大脑不同，人工神经网络在捕捉原始视觉数据时不会局限于只能通过双眼。研究人员进

行了一项实验，把雷达、激光雷达或相机中的数据与视觉系统连接起来，制造出一只比人眼更传神的“人造眼”。此外，还有一些前景诱人的领域等待人们去探索：收集各车辆数据，连接无人驾驶车队的集体学习技术，更好地推动网络视觉感知能力的发展。

目前，深度学习技术已经在无人驾驶领域的几款应用程序中崭露头角。以色列企业移动眼（Mobileye）就在视觉软件中使用类似的算法进行深度计算，并把软件卖给包括特拉斯在内的无人驾驶汽车企业。当谷歌汽车在山景城的街道上审慎移动时，移动眼却正在训练无人驾驶车队识别道路常见物体的集体能力。

新型深度学习算法能分析视频信号流产生的多个框架，也创造出包括动态-深度感知在内的视觉特征。例如，动态深度学习算法能够学会识别一只猫，但不是通过识别猫独有的耳朵和胡子，而是通过识别其躯干的移动特点和猫走路的姿态。应用到驾驶领域，该深度学习网络能通过感知路面物体的移动，得知路面状况：猫不可能突然出现，正常的行人也不可能以50英里每小时的速度飞奔。我们还能进一步地推断出网络具备根据行为方式来识别物体的能力，那些人类看来稀松平常的物体行为方式，往往就是系统识别的线索。

当深度学习软件获得了全新看世界的能力，那么自动汽车对数据的依赖程度就会不亚于汽油之于汽车。无人驾驶汽车的人工传感器需要消耗大量的数据。诚然，它们在消耗数据的同时也在提供数据。当车辆、道路都成为日益丰富的数据来源的一部分，才有可能保障街道安全和交通顺畅，同时也给个人隐私的保护带来全新的挑战。

第十章

燃料——数据

Fueled by Data

无人驾驶汽车既消耗数据也生产数据。驾驶过程变得越来越数据化，我们驾驶的方式将会发生以下几方面的变化：无人驾驶汽车将由高分辨率的数字地图来引导，拥有数字地图资源的公司将参与并维护这新开拓的市场；自动驾驶汽车带来的另一个后果是城市需要管理好交通模式；驾驶的数据化趋势带来新收获的同时，也会带来前所未有的风险。

当汽车完全自动化以后，我们那饱受智能手机与网络泄露威胁的个人隐私，泄露风险将进一步增大。无人驾驶汽车完全可以收集乘客和行人的身份、习惯（甚至详细到他们的目的地和同行的身份）。此外还将面临另一个道德挑战，由于人类社会非常尊重生命的价值，因此无人驾驶汽车软件需要有能力和在交通事故中计算出妥当的处理方式。

我们先谈谈无人驾驶带来的机遇。首先，高清数字地图的广泛使用将有助于一个新兴行业的崛起，同时带来激烈的行业竞争。2015年，诺基亚把旗下的数字地图与导航分公司（Here map）卖给由资金雄厚的德国汽车企

业组成的联盟。这个联盟计划将这些数据应用到本地服务中。

在绘制高精度数字地图的比赛中的领先者是谷歌公司。数十年来，谷歌投资了大量人力和数十亿美金，打造高清先进的数字地图。谷歌公司数字地图的原始数据来源于政府项目，如美国人口普查数据和五十州地理的详细信息。从那时候起，地图信息就被不断更新。新数据由谷歌团队手动添加，谷歌公司派遣专职人员在车内装上激光雷达和摄像头，边行驶边拍摄建筑物、街道标志以及重要的地理信息，找到那些航拍图片中没能发现的隐藏信息。

多年来，谷歌地图项目团队在公司内部被称为“街道真相发现者（Street Truthers）”，他们行驶了500多万英里，收集各种数据。数百名雇员，其中大部分在印度班加罗尔，根据用户提交的错误报告来修正地图信息。这支海外工作团队的作用很关键，他们整理数据，汇集分散的数据库，也收集错误。

高清地图的制作和更新需要耗费大量的时间和人力，由此推测，似乎高清地图的未来市场将会掌握在某些财力雄厚的个人和企业手上。在许多国家，各级政府致力于收集和维持海量的原始数据，记录交通灯的信息或绘制山脉的详细地图，然而其实这些数据中的大部分已不再可靠了，至少在美国情况如此。假如一个人有充足的时间，那么他可以走访各大城市和美国联邦机构，最终收集到可观的数据并建立起数据库。只是如此完成任务效率不高且耗费时间，尤其是那些被收集来的数据需要重新整理，并以一定标准储存后才能调用。

未来的高清地图市场发展多样且利润可观。潜在的客户可能是需要记

录当地街道、路面的交通状况、监控维护路标的使用情况的城市机动车部门。此外，交警部门也想知道哪些路段会导致无人驾驶汽车突然刹车或较为危险。财力雄厚的公司客户则包括保险公司和生产无人驾驶软件的技术公司。

车辆内置的地图信息越详细，车辆安全系数越高，对于潜在客户而言的市场价值就越高。未来的消费者在购买车辆时，会优先考虑那些配置有更加详细的当地街道地图信息的车辆，这样他们就能知道，比方说，在哪一段路段靠边停车比较安全。

高清地图的生产和维护成本很高，但这正是其市场价值所在，也正是许多想跻身该行业的公司要跨过的门槛。然而长期来看，数字地图将会成为常见商品。两股力量将导致这样的转变：

第一，数码相机和深度学习软件性能提升，汽车操作系统的可靠性就从地图数据存储转移到实时场景识别。实际上，当汽车具备视觉智能后，车载地图反而将变得不那么重要。

令人始料不及的是，第二股力量竟然是“自动化”。一旦汽车能自主驾驶，它们就能自动更新并维护车载数字地图的信息。不计人力成本以后，曾经居高不下的地图制作成本将遭遇滑铁卢。庞大的无人驾驶汽车队伍将会成为首屈一指的地图制造者，因为机器人有着过人的精力、耐性和稳定的情绪。

机器人的自我管理

深度学习软件应用在驾驶上产生了一种秒不可言的边际效应。导航软件一开始效率不高，不过只要达到某个临界点之后，导航软件的性能就会持续加速提升。换言之，无人驾驶系统收集的数据越多，它们的驾驶技术就越高明。总结这种良性循环，我们会发现这与机器学习软件的训练数据多少有关。消耗的数据越多，软件的性能就越好，从而推动系统越发积极采集数据。

该良性循环是这样展开的：起初，需要有人类司机来训练汽车学习软件。积累到了一定程度，到达一定临界点后，深度学习软件开始自主导航行驶，同时收集新的训练数据，形成稳定的数据流。新数据用于训练深度学习软件，以提高物体识别的准确率，进一步改进性能。当导航软件的能力提升后，就可以派遣更多的车辆在街道上行驶。当无人驾驶汽车以车队的形式集体工作时，其增长速度会开始以幂次方速度进行增长。车队装配有相机和深度学习软件，将会记录每棵树、每堵墙、每个垃圾桶、每道风景的特色。一天下来，车辆把数据上传到中央数据库，这样别的车辆也能利用这些数据。有了这上千甚至上百万辆无人驾驶汽车贡献自己采集的数据至丰富又鲜活的数据池，深度学习软件的性能才能飞速增长。

车队学习还会以另一种方式加速无人驾驶汽车性能的提高。若是几辆汽车记录了相同的视觉环境，重叠部分将会被交叉检查。由于错误检测软件需要检查多辆汽车，实现数据交叉检查，以确保数据的准确性。假如有

1000辆汽车都报告看见一棵树倒在路旁，并且目击时间一致，观察角度各异，那么确实有一棵树倒在路旁的概率就很高。

深度学习软件的能力最终是由软件自己来提高的，这让人联想到一个有趣的场景，那就是移动机器人有能力提升车载地图的准确率。这个地图构建过程被称为“即时定位与地图构建（simultaneous localization and mapping, SLAM）”。发明这个方法的同学最初命名为“即时地图构建与定位”，字母缩写为SMAL（幸运的是他的导师建议重新调换单词次序，于是一个与著名篮球杂志《SLAM》相同的，响亮的字母缩写SLAM诞生了）。

SLAM本质上是把Bootstrap——一个被广泛应用的Web前端框架——采用在地图上的过程。在这个过程中，移动机器人能很快构建起新的认知环境，以其中一个点作为初始视野点，并由此扩散开来。该过程以机器人选择起点作为开端，接着，它开启视觉传感器记录下该位置所能看到的一切物体，所捕捉到的原始数据被录入空白的原始地图。接下来，机器人逐步推进，并重复扫描，在第二个视野点描绘出周围场景的第二幅地图。先前位置上观察到的大部分物体在当前的视野中也能看到。根据两幅地图的重叠部分，机器人能进行校对，整理成一幅更大型、更精确的地图。对这幅新融合而成的地图进行三角测算处理，根据所看到的物体推断出机器人所在地点以及确定下一步探索的未知地图边界。

由于机器人不断进行即时定位与地图构建过程，它马上就能完整而精确地构建出该区域的局部地图。我们曾看见研究中的机器人使用SLAM方法快速对建筑物周围环境进行学习：机器人以一张空白的地图开始，经过几个小时的游荡，成功生产出一张大型地图。

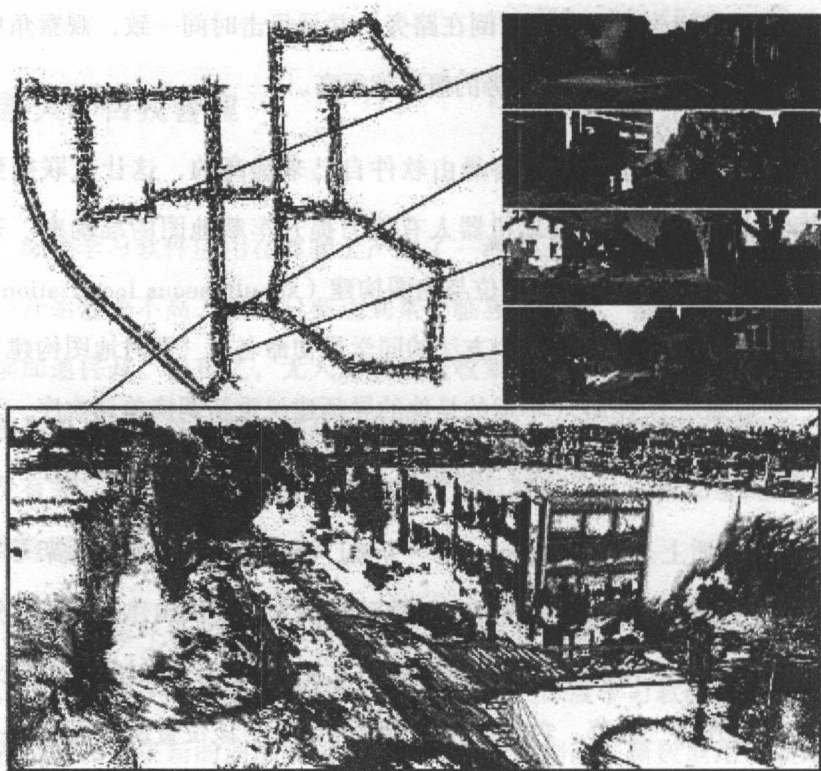


图10.1 使用即时定位与地图构建（SLAM）技术得出的3D地图

来源：Jakob Engel, Jorg Stuckler和Daniel Cremers合著的论文《立体相机的规模直接SLAM技术》（*Large-Scale Direct SLAM with Stereo Cameras*），2015年国际智能机器人和系统大会，1935~1942页。

Andreas Geiger, Philip Lenz和Raquel Urtasun合著的论文《自动驾驶，我们准备好了吗？KITTI视觉基准测试套件》（*Are We Ready for Autonomous Driving? The KITTI Vision Benchmark Suite*），2012年国际计算机视觉与模式识别会议，3354~3361页

SLAM技术赋予机器人自主提升地图准确性的能力。由于移动机器人使用SLAM技术来收集传感信息，机器人储存的地图质量不断攀升。良性循环开始了。当机器人存储的地图精确度与详尽程度提升到了一个新高度，机器人就成了一个更好的导航者，从而能更好地出发，去收集更多的数据来

更新地图。

如今，移动机器人把SLAM技术应用在各个领域，从军事领域到家用扫地机器人。与车队学习相类似，也出现了团队型SLAM。在这个过程中，几名机器人通力合作，把视觉数据进行汇合。SLAM另一个分支技术是3D SLAM，这项技术可以使自动潜水艇绘制出海底地图，使空中无人机描绘出探测过的洞穴的内部情况。

如果有无人驾驶车队使用Booststrap框架来提升高清地图和深度学习的准确性，它们将会采用类似团体型SLAM和3D SLAM的工作过程，反复拍摄每个街景，仔细查看每个路面上的裂痕、颠簸之处和道路标志。当无人驾驶车队获取的图片信息形成平稳的即时数据流并输入软件，就能形成一个不断成长和更新的数据模型，内容将覆盖路上行人的举止行为和该地区路面街道的详细情况。当车队把它们各自驾驶的经验都汇集成机器人认知的合集时，综合起来的经历将等同于一个人类司机几千年积累的驾驶经验。

预测交通情况

当无人驾驶汽车进一步地开展其地图业务，它们的位置、速度和选择路径会生成非常实用的副产品：交通数据。今天我们使用路径规划软件来提供有限的导航来避开交通拥挤。几乎每个人手机上都有一款APP是用来指导我们到达特定的目的地，这些APP较为先进的版本会把交通信息也加入到运算当中。

无人驾驶汽车将包含复杂的交通预测软件，它使用机器学习软件从实时以及历史交通数据中学习。经过学习此前数小时、几周甚至几年的交通模型，机器学习软件将学会识别造成交通拥挤的原因，例如交通事故或是道路维修。其他影响因素包括节假日、学校放学、举行运动会和大型社会活动；天气和冬季的积雪是另一个破坏性极大的因素。

交通预测和计划软件的性能提升曲线与深度学习软件的类似。可用的交通数据越多，交通预测的机器学习模型将变得越准确。一辆车行驶时间越长，它就越了解交通状况——就像经验丰富的老司机遇上交通拥堵时会选择其他路线——交通预测软件的性能越好，就越容易找到最佳路径。

从今往后的几十年内，交通预测软件能够达到我们今天梦寐以求的高水平。当无人驾驶汽车对好几年的交通数据进行分析，我们会发现预测软件揭示了看起来毫不相干的交通状况之间的关系。城市规划者将发现一个交通状况将会间接引起了另一个交通状况，这就是我们熟悉的“蝴蝶效应”。一起不起眼的交通封锁会引起几条相距甚远的道路发生10小时的交通延误。

道路规划与交通预测软件所能做的，远远不止路线规划和指引车辆绕行交通拥堵区域。这种软件将拥有整个城市的交通模式的全景图。然而在某些状况中会用到微观地图，在这种情况下需要在短时间规划出较短路径，规划出到达附近目的地的最佳路线。当无人驾驶汽车综合考虑效率最高的路线时，它能创造出一系列的短期预测模型。例如，如果你在一辆前往杂货店的无人驾驶汽车上，没有指定去哪里购物，这时汽车将会综合附近的交通模式，将你送到能最快抵达的杂货店。

交通系统是科学家口中的典型的“非线性系统（Nonlinear System）”。

它是一种由多种相互依赖和制约的因素，用难以预测的方式作用于彼此的系统。对于许多学科的科学家和工程师来说，在非线性系统中分析这些微妙而复杂的共存关系是一个非常热门的研究领域。股市是另一个复杂系统，很多情况下受到人类投资者的影响，其平衡性会不断偏离中心；热带雨林中的生态系统出现物种波动时也是类型的例子，其中包含了上百种独立或非独立的关系，分别以难以预计的方式彼此影响。

我们都经历过“交通非线性效应”，那就是每次遇上早高峰的时候：8点前出门，你将早一小时到达办公室；晚半小时后出门，你将会迟到。在过早抵达和迟到之间有一个理想的时间点，可是我们很难找准理想的出门时间。早晨的交通用时难以计算，因为能让你9点准时抵达办公室的出门时间不是固定的，它受到许多因素的影响。

交通模式非线性的一个原因是群体行为决定了人们的路线选择。当司机听到交通广播播放交通延误的消息，他会想：“慢着，如果人人都选择了别的路线，那么我继续走交通事故发生的路线岂不是更好？”无人驾驶汽车也会留意到交通延误，但是与人类不同，它们的反应与处理方式更为理性客观。先进的路线规划软件将通过假设场景形成方案，迅速搭建出几种不同的交通状况模型。无人驾驶汽车相互告知所规划的路径，监管软件将根据各车辆情况合理分配不同路线，实现真正的高效出行。

当无人驾驶汽车得到普及后，车辆所采集的交通数据量将持续增长，机器学习的模型质量不断攀升。因此，路径规划软件能提前数小时，甚至几天对复杂的交通状况进行预测和维护。对于未来的政府决策者和汽车制造商而

言，有一个问题悬而未决，那就是该如何定义交通预测软件的精确度。

载人的无人驾驶汽车并不是高效的路径规划软件交通预测软件的唯一受益者。商务货运以及快递行业也是这当中的受益者。为商业配送车辆进行路线规划可比家用轿车的难度大得多。受以下几项物流因素的影响，路线规划的难度大大增加。

每天穿梭于城市内家家户户的自动配送车队需要不时在公路停车进行配送服务，大量的停车需求使路线规划变得棘手繁杂，绝对可以媲美气象预测的高难度。个人汽车的路程较短，通常少于30英里。而配送车辆可能持续驾驶一整天。因此，路径规划软件在规划时，必须在一整天或一整个星期的路线中考虑到多个停车点的存在。

然而，复杂的事情还在后头。如果无人驾驶配送车需要在好几处不同地点多次停车，会产生将近5亿种可能的路线排序。这项工作的难度在于卡车和配送车辆进行路径规划时，受“有时间窗的车辆路径问题（vehicle routing problems with time windows, VRPTW）”的影响非常大。如何计算出理想的配送次序方案是复杂系统研究中的一个经典操作顺序问题。最佳路线不仅指距离最短的路径，还要求顾及当天配送车辆最佳的停车卸货次序安排。

路径规划机器学习软件将成为驾驶经验丰富的司机。当无人驾驶汽车和无人驾驶卡车普及后，高清地图信息将完整覆盖路面的每个细节，包括准确的地点、状态和变化情况。而随着路面详细情况结合全景智能路径规划分析，无人驾驶汽车的水平甚至还能更上一层楼。

由于人类司机不在考虑范围之内，那么道路上也不需要实体标志，它将被汽车3D模型中的虚拟“标记”替代。无人驾驶汽车不需要查看道路标志也能清楚知道限速的要求。软件将持续追踪危险地段的情况，比方说，遇上气温急剧下降，某座桥面冰封路滑时。当无人驾驶汽车被普及后，今天单一的限速将会被灵活的、适应性高的速度限制所取代。新型的速度限制将考虑道路和交通状况因素，同时也考虑照顾一些无人驾驶车辆的特别性能。

当大部分的道路车辆都走上全自动化时，无人驾驶汽车的一些优势才会真正显现出来。其中的一个优势就是自动化“交通优先权”系统的实现，它能在紧急救护的需求下暂缓交通压力，把需要急救的乘客尽快送往医院。由于人类司机不完全自律，也没有交通全景的大视野，不会完全按照所规划的路线行驶，所以如此环环相扣的交通优先等级权，必须要求所有的车辆都已经全自动化。

自动交通优先权系统的优先等级可能会划分好几种不同的优先等级。最高优先级别模式当然是急救车辆，或是车上有急需帮助的乘客。其次是监护人不在身边的儿童。下一个级别是去上班或是日常出行的成年人。最后的等级是商用配送车辆，它们被归为速度较慢、效率要求不高的类型，应让道给更高级别的车辆。然而这个系统的潜在争议问题是城市规划者可能设立会“有偿优先市场”，对交通优先权利进行买卖，情形类似于当某航班的预定超额时，航空公司会鼓励购买折扣票价的乘客搭乘其他航班，给他们补偿退款。

个人隐私

在无人驾驶中，另一个亟待解决的常见问题就是无人驾驶车辆的乘客隐私问题。智能手机与社交媒体已经产生了许多包括个人隐私在内的道德挑战问题。无人驾驶汽车也会带来类似的问题。理想情况下，消费者监管机构将会将对人们的隐私进行保护，就像此前避免政府部门以监管之名或数据公司中介侵犯个人隐私那样，机构将进一步把管理范围延伸至无人驾驶汽车。

然而无人驾驶汽车将面临较为特别的隐私问题。车子是移动的，一旦被装配上高分辨率的视频摄像头，外加远超人类的感知与识别能力，汽车可能成为“机器人间谍”。滥用这种能力的可能性很高。车子能拍下乘客或路边行人的照片，进而传输给人脸识别系统软件。软件能将收集来的人脸数据传递给政府。危险性降低了，但烦恼却增多了。当各大公司的市场部急需人们的衣着、出行方式和其他习惯的数据，用来研究个年龄层的穿衣习惯或喜欢就餐的地点时，汽车软件能收集并发送相关的大数据。

行人和车内乘客都需要一定的隐私保护，或是公开透明的数据管理规则。每位无人驾驶车主都应当指导车子收集了什么数据，以及谁有权使用这些数据。假如无人驾驶汽车软件制造商打算把乘客数据卖给第三方，客户应当有权决定自己的个人数据是否用于商业用途。

如果未来的个人出行主要来自于无人驾驶车辆，管理车队的企业需要

严格遵守个人隐私保护指引。由于无人驾驶出租车的乘客量大、路线长，对于数据交易员来说既像是一座宝贵的“金矿”，又仿佛打破了政府监管，收集了大量窃听素材。乘客应当能够基于个人意愿掌控无人驾驶汽车收集的自己的数据的流通情况。在减免车费的前提下，也许有乘客愿意完全共享数据。

隐私还涉及另一项安全问题，那就是软件的脆弱性。任何软件操作系统都容易受到错误操作的影响和恶意入侵的攻击。正如我们在前几章讨论的，软件硬件都有一些潜在的安全漏洞。汽车的CAN总线协议对于潜在的黑客来说，就是一处没有特别防卫的入口。从软件层次来说，无人驾驶汽车的操作系统需要经过特别设计适当添加冗余设计。一旦发生稳定性问题，备用系统就能马上启动。

如果车辆接收外部数据，不管出于何种目的的数据转换延迟都可能造成安全问题。仿冒GPS定位信息就会成为新型犯罪手段，高清数据地图也将成为下一个犯罪目标。提供地图数据传输的渠道必须确保安全性，数据源的真实性需要进行甄别鉴定。路线规划与交通预测软件把车辆目的地的数据发送到更大型的数据系统中，并时刻保持消息灵通，这里面也包含了潜在的安全问题。

道德伦理规范

数据隐私与安全性是由无人驾驶汽车产生的、引起重视的道德挑战，另

一个道德挑战是汽车对紧急情况的回应。当面对突发意外，无人驾驶汽车使用数据和软件，而不是人类的直觉，来计算出最佳处理方法。为了赋予车辆合宜的反应能力，程序员艰难地通过量化数值的途径来保障生命财产。

当通过系统软件开车时，人类司机不再需要在车祸发生瞬间，作出“留哪方性命”的选择。这时经过程序员的提前预设，软件的决策逻辑将发挥作用。经过特别设计，软件能针对事故作出妥当回应，把破坏程度降到最低。想要制造这样的软件，我们人类社会必须首先定义好什么是“理想的处理方式”。在此过程中，就涉及人类如何判断生命与财物的价值。

给生命与财产赋予价值的做法提出一个让人不太舒服的问题：如何界定价值？谁来界定？车内的乘客有既定的价值，路上的行人也有。车祸中出现的可接受的处理结果包括，财物损毁或车辆损坏。如果无人驾驶汽车在编程时被写入的指令是“只将财产破坏降到最低”时，我们难以想象这样的指令执行起来会是何种情形！

当我在演讲中与观众谈论无人驾驶汽车的价值与潜在影响时，总会有人举手提出以下问题：“即将发生致命事故时，无人驾驶汽车是觉得撞向路旁的两名婴儿呢，还是牺牲车内的五名成人呢？”这类道德选择问题还将会在新的情况中被再次提出，但是核心问题出自经典的“电车难题（Trolley Problem）”，一道在过去几十年中的哲学课堂上被热议的问题。

1967年，电车难题由英国哲学家菲利帕·福特（Philippa Foot）提出，是一道描述与道德相关的难题：电车司机只能沿着轨道行驶电车，轨道的一端有五个人在工作，另一端有一个人，无论电车开向哪一端都会撞死相应的人。大部分读者经过简单计算，认为五条人命比一条人命有价值，觉

得这道题目非常简单。但电车难题之所以变得更复杂，是因为其他畸形的选择最终将会导致情形选入两难的境地。

电车难题并不只是针对无人驾驶汽车。最近，在纽约北部的伊萨卡中心城区，我们亲眼目睹了一出无异于电车难题的悲剧。一辆卡车从陡峭的山坡滑向伊萨卡中心区，司机已经将刹车踩到底了，然而车辆并没有停下的趋势。司机将面临着痛苦的抉择，这两吨载重量的失控卡车将撞向何方？最终司机避开了一群建筑工人，选择将卡车撞向附近的咖啡厅，27岁的阿曼达·布什（Amada Bush）当场身亡，她是一位年轻妈妈，正利用夏天这个周五下午担任侍应生赚些外快。

一位学者把大众关心的道德问题描述如下：“在车祸中，自动汽车的主要弊端是不同于人类司机可以即时思考如何处理碰撞，自动汽车的决策来源于程序员提前设置好的程序。”面对这样的表述，我们不禁会问，为何人类凭直觉作出的碰撞决定就可以被接受，而汽车的碰撞处理由于提前被设置好，就成了道德问题？第一，无人驾驶汽车的驾驶更为理性，能够快速进行风险和利益分析，这比一个自私、疲惫的醉汉的决策要好得多。第二，无人驾驶拥有360度的感知器，掌握的信息也更全面。

在我们看来，无人驾驶汽车的设想中的道德问题并不是基于“驾驶过程中，计算风险和价值哪个优先”的事实。所谓的道德问题实际上是因为这些计算是由人工智能得出的结果。最关键的问题并不是无人驾驶车辆是否有道德，而是预先设置的车祸处理方式的逻辑是怎样的。

面对危险时，每个人心里都会衡量风险与收益的关系。有人认为最好的方法是即使代价是牺牲几个人的生命，车辆也应当尽量避免对司机的伤

害；另一些人的反应可能是避开所有人群，牺牲自己与车子。

大部分有经验的司机知道，在驾驶过程中，他们持续地作出一连串涉及衡量性命财产价值相关的决定。整个过程通常是无意识的。我们当中的有人可能不止一次要作出紧急的决定，避免撞向小生命，比如一只横穿马路的松鼠。我那理性的大脑不会发问道：“这样做是否值得？”事实上，我使车子转向，避开松鼠的举动表明我内心是计算过的。松鼠的生命比起车子猛然转向可能引起失控和事故的小概率风险，更有价值。

我的风险价值计算有时也能变得十分复杂。如果我的车后排坐着三个孩子，我可能不会选择忽然转向，决定测试一下松鼠的运气。如果当时的道路结冰路滑，行人拥挤，我的道德选择很有可能把车子内外的人的生命也考虑进去，价值显然比松鼠更高。这样，我就不会进行转向。如果假想的松鼠被换成一只狗，或是一位拄着拐的老人，等等，司机的计算可能会更复杂一些。

我们当中的有一部分幸运儿，暂时还没有遇到重大交通事故，没有被迫当众作出如此痛苦的抉择，也没有在抉择中展现出我们在处理不可避免的交通悲剧时的反应方式。由于无人驾驶汽车迫使我们公开作出这样痛苦的抉择，便激起了一阵惶恐。更具挑战性的是，在面对紧急情况时，无人驾驶需要我们把人类所认同的道德规则用于指引系统作出决策。

无人驾驶汽车需要我们认同一系列用于指导车辆紧急决策系统的道德准则。在公平民主的社会中，这套道德准则被大众认同，无人汽车制造企业也自觉遵守。不仅是应对悲剧发生的决策应当得到大众认同，当灾难发生时，汽车灾难的指导手册应当是公开透明，经得起事故考验的。类似于

空难发生后航管局官员采取的步骤，在发生车祸后，无人驾驶汽车的指导手册“黑盒子”应当受到保险专业人士和执法部门人员的检查，他们最终会分析软件当时都作出了怎样的决策，采取了哪些行动。

当这些道德准备变成了法律，结果将产生新的道德滑坡和犯罪。想象一下，假如有黑盒子显示，汽车公司把车子的价值计算原则设置成车内人员的生命价值高，车外人的生命价值为零；或者车祸后，黑盒子记录显示，因为预先写入程序时的疏忽，软件被设置为“最小化车子的损坏程度，而不顾其他”。

自动驾驶汽车将会挑战我们的价值观和可靠性，并把驾驶从凭直觉转换到单纯依靠数据和逻辑。随着无人驾驶技术的普及，城市的设计也将会经历另一种变化，那时候停车场都消失了，车子被数据驱动制定最佳路线，缓解了出行难题；自动交通运输将给予我们的工作和生活地点更多选择；一些职业会消失，当人类司机从企业的求职表中被拿掉以后，新的商业模式将会出现，从而催生出新的职业。

第十一章

涟漪效应

The Ripple Effects

至此，本书已接近尾声，我们还想提出一个问题，如果这个问题没有解决，那么任何对于无人驾驶的深入探讨都将是不完整的，那就是媒体炒作。近年来，无人驾驶汽车成为媒体大肆宣传的重点。“解放双手，解放双脚”的无人驾驶梦想之路精彩而漫长，然而，我们应当从中汲取教训：即使人们对自动驾驶汽车渴望已久，但人类的欲望是无法保证自动驾驶汽车真正出现的。

技术行业内的资深人员也许还能回想起，被称为“赛格威（Segway）”的两轮体感平衡车。2001年这种平衡车进行了大张旗鼓的宣传。就在宣传开始的前几天，《时代周刊》上的一篇文章解读了平衡车的神秘发展过程。文中提及史蒂夫·乔布斯在一本技术类书籍中，预言平衡车会像电脑一样普及；风投界传奇人物约翰·杜尔仔细思量“平衡车未来的影响力是否将会比互联网更大”。事实证明，他们都错了。今天，这种平衡车仅仅成了给仓库员、邮递员和商场保安提供代步服务的小型短途的交通设备。

诸如平衡车之类的科技层出不穷又迅速衰落，不禁使我们产生了一个有趣的疑问：“为什么同是前景乐观的新科技，其中一些能够对部分行业产生毁灭性的打击，而另一些却与宣传预期相去甚远？”与其他人一样，我们用了很长时间来思索该问题。经过多年时间，我们研发出一套用于评估新技术的长期潜力的小型而实用的测试工具，叫作“零原则（Zero Principle）”。

以下是零原则的工作原理。撼动传统行业的新兴科技都有一个共同特点：产品推出后，一项或多项生产工作的成本将降到几乎为零。实际上，遵循零原则的技术推出几年后，产生了极大的行业影响，最终成为行业革命的导火索。

我们来看一些历史先例，从蒸汽机开始。如果你是18世纪末英国的技术观察家，你会考虑对新兴的商用蒸汽机进行投资吗？如果你用零原则来衡量这个新产品的投资价值，你能马上看到它的潜力。

蒸汽机大幅降低了动力行业的机械成本。在发明蒸汽动力以前，工厂和磨坊的动力来源都受到直接、间接的现实操作限制。工厂的动力来源于水力，因为选址必须靠近湍急的河流瀑布，依靠动物作为动力同样受到区域限制，而且动物很快觉得疲惫，需要额外的照顾和喂养。

当商用蒸汽机被引入工业时，直接和间接的动力机械成本都被“蒸发”掉了，转变了制造流程，最终引发工业革命。蒸汽动力降低了钢铁产业的成本，并引发了下游行业的创新。价廉物美的钢铁催生了好几项下游行业的诞生，比如铁路运输和能抵御炮火的装甲舰制造业。

将近200年后，另一项产生深远影响的发明——计算机出现了。和蒸汽机一样，计算机势必会影响整个行业的发展，因为它降低了曾经十分高昂的计算成本。

在人类历史上，计算过程曾经一度缓慢、昂贵、依赖人力又容易出错。即便是经过专业训练的计算专家，使用最好的计算工具，每小时也只能运算几百次。20世纪50年代，计算技术发展成熟，计算成本开始大幅下跌。随后的几十年间，硅芯片替代了分析技术，提高计算速度和稳定性的同时，还降低了成本。小企业也能负担得起原来昂贵的计算，而且计算结果又快又准确，计算机器孜孜不倦。到了20世纪末，平价的计算机催生了软件行业和电子游戏行业的全球化市场，普通人也能购买个人电脑和游戏机。

历史显示，这两种不同的技术——蒸汽机和计算机——有着共同的潜在特点。它们的出现去除了主要的工业成本障碍，改变了许多不同行业领域的商业实践，快速影响着人们的工作和生活。今天无人驾驶技术的发展也可能会出现颠覆性的局面。只有时间才能证明我们是否即将步入另一个社会大变革时期，或是另一个过度宣传却毫无成效的新兴技术。

让我们把零原则应用在无人驾驶的案例中，看看哪些直接或间接成本能减少。无人驾驶汽车所能减少的最大经济、社会成本之一就是交通事故造成的伤害。另外，能节省下驾驶的时间。对于普通人来说，花在驾驶上的时间是间接的机会成本。对于运输公司而言，人类司机的时间成本直接以薪酬的形式体现，是决定货物运输过程的主要因素。最后，由于无人驾驶技术去掉了人为易发生事故的因素影响，汽车和货运车辆不再需要以“体型庞大”和“专为安全性而设计”的形象出现，同时无人驾驶还能降

低燃油的消耗，催生丰富多样的车身和车型。

如果我们仔细梳理一下，会发现无人驾驶这种自动汽车技术将四项核心成本减至接近零：

1. 接近零伤害。驾驶是高风险行为，无人驾驶汽车能极大降低车祸的直接或间接成本，也减少了每年因交通事故住院的医疗成本（美国全年的医疗费用在180亿美元），以及相关的工薪损失（每年330亿美元）。医疗、保险、器官捐赠等收入与交通事故挂钩的行业将丧失部分收入来源。

2. 接近零技术。无人驾驶汽车去除了一项客运或货运的成本：工资。货车司机的工资是运输货物、商品的一项重要成本。乘坐出租车的成本也和出租司机的工资有关。

3. 接近零耗时。无人驾驶汽车将驾驶的时间减少至零。平均每天美国人花在驾驶上的时间是3小时，每年耗在交通拥堵上的时间是63小时。以前用于驾驶的时间可以用于工作或个人生活。

4. 接近零尺寸。有人驾驶的车辆处于安全设计的限制，体型庞大沉重。无人驾驶汽车的事故可能性较小，因此车型更小，重量更轻。没有司机的货运车只需要与所运输的物品大小一致即可。

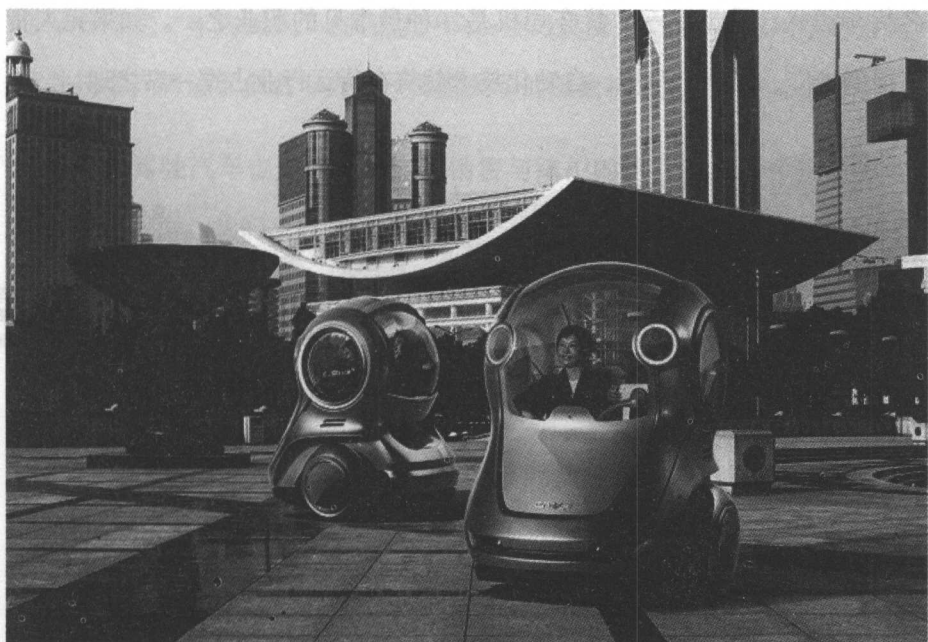


图11.1 通用汽车的电子概念车EN-V，是通用汽车公司与赛格威共同研发一款两人座的自动汽车

来源：通用汽车公司

工作岗位

人类驾驶的直接或间接成本决定了将近一个世纪以来的商业模式。由于无人驾驶技术能减少甚至直接去掉这些成本，最终影响了部分行业的生存，导致某些行业门店关张，旧的工作岗位消失，新的行业和岗位将陆续涌现。

首当其冲的是货车驾驶。这份稳定、高薪的蓝领工作曾大幅抵御了离岸外包和自动化技术的冲击。根据2010年全美普查数据，美国有将近350万

名货车司机。数据显示，货车司机是29种最常见的职业之一。如果无人货车得到推广，用不了几年，自动化技术就将在货运行业占据一席之地。



图11.2 货车司机成为2014年美国绝大部分州最常见的职业
来源：美国国家公共广播（NPR）

货车司机不是唯一受到无人驾驶冲击的工作。出租车司机和其他驾驶员也会发现自己处于失业的边缘。专职的出租车司机目前已经受到了越来越普及的Uber、Lyft等共享租车服务的威胁。在那种模式中，任何拥有车辆的人都能当司机。无人驾驶出租车将为全美国233,700名出租车司机的职业生涯敲响丧钟。

Uber的CEO，特拉维斯·卡兰尼克（Travis Kalanick）认为出租车运营成本中最大的一项就是司机的工资。在一次论坛中，卡兰尼克说：“如果车内不需要司机，那么在任何一个地方叫Uber的费用都会比自己买车更划算。”为

了早日实现“车内无人”的理想，Uber公司投资了550万美元用于研发无人驾驶技术，聘请了多名卡内基·梅隆大学的国家机器人工程中心专家。

无人驾驶汽车也促使支撑汽车销售与维护的巨大经济价值链中的其他岗位发生转型。我们一般从汽车经销商手中购买汽车，在服务中心的等候区内，我们用一次性塑料杯子喝着免费咖啡，随手翻阅厚厚的旧杂志，这时隐约听到雇员对于下周工作排班的争论：两位雇员正在争论谁应该上周末班。他们不知道未来的一二十年间，星期六将可能变成他们的常规工作日。

当汽车能够自动驾驶时，人们不再需要在工作日来汽车经销中心，无人驾驶汽车将会自行开到店里来。由于大部分人都担心车子检修会给自己的上班出行造成不便，他们会在凌晨3点下订单，让车子早晨回到家门口。至此，美国739,900名目前就职于汽车经销中心的技术员也许会为谁上凌晨3点的班而争论，但听到他们的声音的将只有车，没有人。

多种职业消失了，那么是否会有新的替代职业出现呢？目前尚无定论。20世纪40年代，经济学家约瑟夫·熊彼特（Joseph Schumpeter）提出了“创造性破坏理论（Creative Destruction）”，用于描述因为某种技术的出现摧毁了原有行业后发展出的重建过程。这个重建过程触及经济的几个主要部分，包括设备使用和管理结构。其中最具争议性和影响最大的表述当数对工作岗位的破坏。

熊彼特主义的经济学家积极看待创造性破坏的循环过程。他们相信破坏性技术能够创造出更多更好的职业，即便破坏性技术暂时取代了一整个行业的工人，但长期作业下将会有更多更好的职业被创造出来，因为从旧行业的灰烬中将诞生出新行业。

创造性破坏的支持者指出，无数的创造性破坏案例是时代发展的产物。桌面印刷软件淘汰了原来铅字排版的工作，但它创造出更大的市场，那些富有创造力的排版从业者有了自己的工具，设计和印刷自己的小册子、书籍和通信刊物。以Expedia.com为代表的新兴网站几乎让传统的旅行社消失殆尽，同时却也点燃了更大、更活跃的全球旅行市场的热情。

尽管人工智能软件和机器人日益精细复杂，功能成熟，但创造性破坏催生更新更好职业的观念却不太能经得起考验。在上两个世纪，被新技术所取代的工人，至少从理论上讲，能够在新技术行业中找到工作。不过近年来，随着工作自动化趋势从工厂的操作工攀升到白领的分析型岗位，使得工人重新就业难度更大了。由于现代信息技术提升了工作效率，在旧岗位消失的同时，虽然有大量的新工作出现，可是规模小，薪酬低。

在人工智能与机器人能够承担越来越多的工作的今天，有两个关键问题决定了创造性破坏真的有创造性，还是仅仅起到长期破坏作用而已。问题一，新工作是否比旧工作要好，是否更安全、更有趣、薪酬更高？问题二，该工作被替代的周期长度，即被替代的工人失业仅仅维持几个月，还是长达好几年？理想的情况下，被机器取代的工人经过重新培训，并很快上岗。然而，在最坏的情况下，这些工人将多年被排斥在劳动力市场外。

机器夺走了人类的饭碗，这个潜在的破坏作用将进一步恶化收入差距这个日益严峻的全球性问题。反对贫困的慈善组织乐施会公布的研究结果显示，富人越来越富有。2009年，最富裕的1%人群拥有全球财富的44%；2014年，这个数字增长至48%；在未来几年还将继续增长，最富裕的1%人群将占有全球过半的资源与财富。

对于创造性破坏的深度分析不在本书的讨论范围内。我们在此引入这个概念为了说明无人驾驶汽车的出现——像在此之前的破坏性技术一样——重塑行业结构，促使上百万个岗位消失。最坏的打算就是，无人驾驶汽车毁灭多个行业，削减了人们对司机的需求，加剧正在扩大的贫富差距。另一方面，较为乐观的结果是无人驾驶汽车能够带来经济的增长，为社会各阶层人们创造高质量的工作。

交通事故带来的工作收入

有一句古老的英语谚语说，一个人的收获来自另一个人的损失。这样看起来，无人驾驶汽车能够挽救生命、节省燃料和时间，那么受损失的就是部分商业模式与交通事故相关的企业。这类行业的范围很广，从汽车保险到个人伤害维权的律师，从实体店到零件供应商，从高速公路的巡查人员到预防性驾驶课的老师，从器官捐赠机构到急救室医护人员，从交通法庭到监狱……

美国人每年因交通事故住院的时间累计多达100万天；美国约有20%的器官捐赠是来自于交通事故中丧生的人；美国监狱有相当一部分囚犯与人类的驾驶弱点有关——1997年，7%的监狱犯人和14%的缓刑犯人狱原因是醉驾。

美国每年汽车相关的各种问题产生的罚款多达几十亿美元。一个让人不悦的真相是，因为各项法律与停车强制执行细则开出的罚单，不是基

于维护公共安全，而是为了维持政府的资金。如果人类不再超速或违规停车，政府的许多重要收入来源就消失了。仅在美国，每年因超速开出的罚单收取的罚款达到60亿美元。平均每年，纽约市因为违规停车的收入超过了6亿美元。当全世界都是遵守交规，还自带数据记录的无人驾驶车辆，各城市与各州所依靠的收入来源也随之不复存在了。

汽车保险业在无人驾驶的背景下也需要转型。年收入2000亿美元的汽车保险行业将会面对何种程度的影响，现在还众说纷纭。一方面，保险公司的利润可能增加。因为无人驾驶汽车的出现，事故费用索赔少了。另一方面，由于车祸发生概率大大降低，无人驾驶汽车车主可能会要求保险公司降低保险费率。在美国，汽车强制保险是由各州政府管理的，消费者能向州政府施压，甚至要求取消汽车强制保险。

无人驾驶汽车迫使法律和保险公司重新考虑交通事故中的过错分配。保险法专家罗伯特·皮特逊（Robert Peterson）写道，保险与侵权法（认定过错归属哪一方的法律）就像一对异卵双胞胎：并非一模一样，却是彼此的映射。皮特逊还说无人驾驶汽车普及后，伤害的责任很可能绕过司机，归于销售商和制造商。如果责任的重担绕过人类司机，落到无人驾驶汽车的头上，保险公司将不得不改变赔偿成本的结构。传统的汽车保险费取决于司机的个人风险。如果司机不再牵涉保险相关的计算费用，传统的司机风险评估因素，例如年龄和性别将不再采用。

如果无人驾驶汽车以产品的方式来投保，将会评估汽车生产商在交通事故中应承担的责任。这样，无人驾驶汽车的保险费用评估，是把车辆视为产品来评估潜在的风险。新的汽车风险评估方式还有待制订，可能涉及

我们之前讨论的汽车操作系统中的人身安全评级，或者无人驾驶车辆累积的深度学习里程总数。

如果保险公司把乘客安全责任归于汽车制造商，那么在交通事故中，哪一方的汽车制造商来承担过错呢？这类合法性说明的工作量很大，而且需要落实。两个主要的汽车制造商：分别是提供汽车操作系统的一方与提供汽车机械车身的一方。由于无人驾驶汽车操作系统的复杂性，软件与硬件紧密结合，相互作用，因此判定最终过错方是一个艰难的过程，尤其是当事故中涉及错误操作或意外的道路状况的情况时。

另一个需要进一步探索责任划分的领域是汽车维护。比如当一台无人驾驶汽车或出租车不能正常更新车载的软硬件时，责任就将不再清晰。有公司的律师可能会辩驳道，如果有维修人员篡改了操作系统或破坏传感器等重要硬件，那么将是这位维修人员，而不是制造商或车主，承担事故的责任并损失费用。

问题的解决办法之一，是汽车保险公司会为出租车公司提供“每一次打车”保险作为汽车软件的保证金。乘客在乘坐无人驾驶车辆前，是否同意接受乘坐该车可能造成的伤害性后果，这就像今天我们购买电脑的时候，有一张需要点击确认的许可书一样。也许有一天，在我们乘坐无人驾驶出租车之前，还需要签署多达200页的许可书，以表示我们是否同意放弃索取补偿的权利。

新型车身

我们预计，无人驾驶汽车的里里外外看起来都会和现在的汽车不一样。方向盘会消失；仪表盘将变成灵活的工作区间；车厢中将配备人们休闲娱乐所需要的设备；车子外部不再需要侧镜或尾灯。

正如我们在第二章所讲的，未来汽车行业将划分为两大类：制造实用标准化运输工具的公司和特殊顾客用途车辆的制造商。特殊顾客用途包括一人小型车辆到能在车内休息或工作的大型豪华车辆。大部分的车辆都会被卖到交通公司。在特别用途汽车的市场上，个人汽车体型更小，需要经过特别的汽车设计。

无人驾驶汽车带来的积极作用之一就是消费型汽车的设计将再次兴起，这是自动化汽车的新黄金时代。20世纪50年代至60年代，汽车设计师创造出巨大炫目的汽车尾鳍，漆上高档缤纷的颜色。无人驾驶汽车将专门设计豪华或灵活设计的多用途内饰，为人们在车内休息、饮食和工作创造条件。

当安全性的关注度降低，客户可以使用自动设计软件来定制不同的车身。即使新车的购买者对汽车设计了解甚少，也能在线浏览虚拟展示作为参考。如果他们的目光停留在某一件漂亮的设计上，计算机将留意到他们瞳孔的扩大，为他们展示这款设计额外的迭代产品。当然，消费者的汽车设计应当受到空气动力学定律的约束。考虑到“街头行驶的合法性”，消费者的汽车设计也应当符合基本的安全准则才能获批生产。

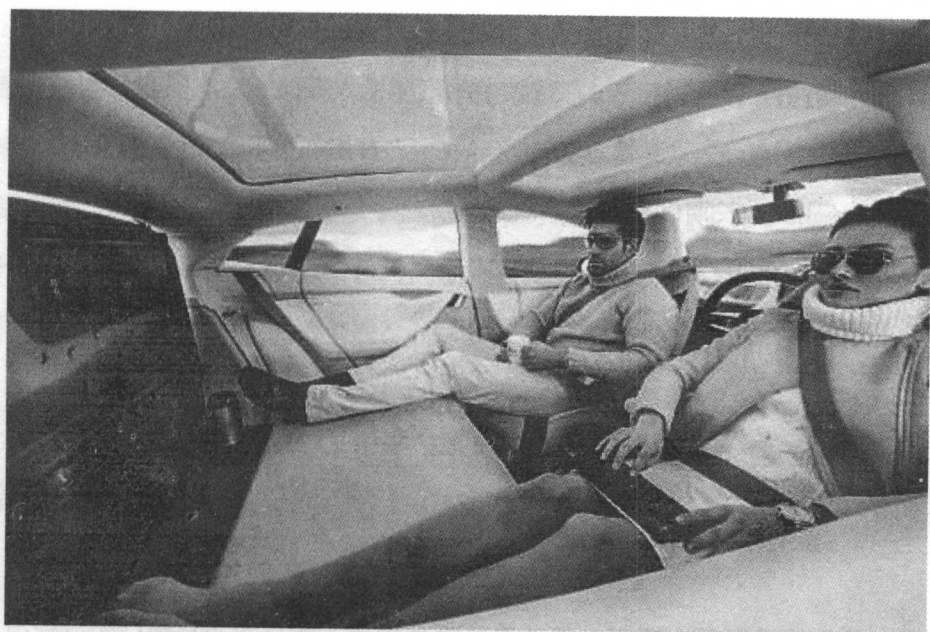


图11.3 乘客在无人驾驶概念车内玩电子游戏

来源：林斯比得汽车公司

买方与软件以互动的形式推进设计过程，反复几次修改新车的车型和风格。一旦买方的设计方案最终确定，制造商需要用一周时间去落实。车身嵌板与底盘需要用碳纤维3D打印出来，创造出宛若鸟骨般又轻又结实的框架。3D打印完毕的车辆开始接受编程、测试，一个小时后到达车主的家。到达目的地后，新车能马上使用，打开车门，向新车主和乘客进行展示。

在无人驾驶汽车时代，一些自动化的工作将变得非常精细化，一份高薪的新工作即将出现，那就是软件机械师（software mechanic）。软件机械师是自动工程师的一种，专注于汽车操作系统的某个领域，例如底层、中层和上层控制系统。软件机械师能根据不同汽车的人身安全评级为顾客提供专业的购买意见。

另一类的专业岗位是噪声扰动的降低。人们如果尝试过在休闲房车的床上入睡，就等同于体验过在行驶中的汽车里入睡的不适。隔音（Acoustic Dampening）、震动消除（Vibration Cancellation）和动态补偿（Motion Compensation）长期以来都是多个工程领域的专业技术。预测信号处理方面的专业自动化工程师（就是头戴式降噪耳机的那种技术）能使乘坐无人驾驶车辆的体验更稳定、更宁静。

新的市场

我们预见的另一项无人驾驶汽车附带作用，就是在产品推广的过程中赢得新的物理维度。无人驾驶软件将会成为新的市场推广“必争之地”：路线投标。传统的地图指示标出“地标兴趣点”，包括餐馆、风景名胜、停车场、购物中心、博物馆，等等。未来的无人驾驶世界中，一些公司会付费给制作高清地图的公司，要求将企业被标注成为地图上的“地标推荐”。即当乘客乘坐一辆无人驾驶汽车，并告之目的地，这时车辆的操作系统会告之，根据他的目的地，为他推荐一些同类型的商家，并询问他是否愿意逐家停下看看。愿意接受推荐的乘客在沿路的每家商店购物将可享受一定额度的折扣。

商家用返还或打折的手段提升顾客体验的做法并不新鲜，每天人们都能在免费网站上阅读到各种在线广告。我们中的大部分人已经接受了这种“免费”的交易，我们默许商家往电子邮箱和文本信息中投放各种广告，以换取免费使用的邮箱和平价的电话服务。

无人驾驶汽车的乘客也会对这些条款讨价还价争取自己的权利。精明的乘客知道该如何用减免车费，或是为自己的车辆换取更便宜的油费作为筹码来交易自己的到访。餐馆会同意为部分外出聚餐的家庭分摊油费，当然前提是这些“每次外出就餐”，车辆系统记录显示他们在出行上花了超过200美元油费的前提下；部分客人会以时间量来做交易，如果商家希望他们到访路程较长，路况不太方便的消费地点，顾客将会得到较低的打车费用。

新型的零售业

精心规划顾客路线的餐馆和企业将会从无人驾驶汽车中受益良多。另一个即将发生改变的行业是零售业。这个行业在上个世纪已经经历好几次转型，导致转型的影响因素首先是规模化生产，接着是大购物中心的出现，再接下来是折扣店，最近的一次是网络商店。无人驾驶汽车将会引发零售行业的最后一次变革。

在古代，珍贵的香料经由骆驼从中东运输至欧洲；今天，油罐车运送到世界各地的港口，在那里等候的集装箱把货物装载上卡车，运送至装卸码头。无论货物的性质如何，交通运输费都占据了产品价格的一部分。对于实体货物而言，运输费用占到农产品价格的14%、工业制成品的9%。

买方需要承担的货物运输费用一直以来都是零售行业决策的关键因素。现在大公司和规模经济才能享用的主要竞争优势已经被无人驾驶汽车打破了。曾经一度高昂的产品运输费用下降了，小企业也能在价格上与大型企业竞争一番。



图11.4 自动运送方式

来源：Starship Technologies

面临破坏风险的零售业之一，是本地出产的农作物和食物的零售业。消费者非常乐意掏钱购买新鲜的、本地出产的、从农场直送到餐桌的食物，消费者也赞成发展好几处小型区域农场，带动农业经济的多样化。不过尽管当地出产的食物有着独特的吸引力，但大部分人仍然会在杂货店里购买工厂化农场出产的食物。由于工厂大批量种植和运输食物，工业规模的合作农场能够以较低的价格来售卖产品，因此消费者能够在连锁商超里买到较低价格的食物。

无人驾驶汽车为小型的食物作坊与集团工厂化农场开展竞争创造了条件。我们的一位朋友在纽约上州开了一家绵羊牧场，他们在郊区种植有机蔬菜、牧养温顺的绵羊，过着平静的生活。他们的市场化策略中成本最

高、最烦琐的部分就是每两周一次往返于城市农产品市场，去售卖新鲜的农产品。每两周一次在凌晨1点从牧场出发，到纽约市的车程需要4小时，他们休息1个小时后，要花上2个小时做准备。经过集市上漫长的一天后，深夜开车回来。假如他们能购买一辆无人驾驶汽车，那么原本耗在路上的时间便能在车上得到充分休息。更棒的是，无人驾驶汽车有助于拓展他们的商品影响力。他们可以把部分商品装载在无人驾驶车辆上，运送至其他区域的集市，并在当地安排一位销售员负责销售，根据销售量支付他佣金。

部分区域制造商的商业模式包含了繁重的运输费用，而小型农场只是这部分的制造商中的百万分之一。任何生产低利润的生鲜产品的公司都需要高效的配送网络来维持利润。由高效的无人驾驶运输车辆组成的网络能使小企业开出极具竞争性的价格，与批量生产的企业商品一较高下——货运成本的主要因素之一是司机的薪水。公司之所以使用一辆大货车而不是两三辆小型汽车把货物运送至零售商店，是因为这样才无须雇用多名薪水高昂的货运司机。对于费用谨慎考虑的运输公司会仔细地规划路线，并且把个体货运量积攒成巨大的集装箱。这种方式的使用产生了被广泛采用但效率不高的“辐射状”运输模式，在该运输模式下不仅配送效率低，还经常使车辆多走“冤枉路”。详细如下：单个货物需要首先到达最近的集中点或配送中心，当集中点或配送中心的货品积累到一定程度，再被运往目的地。

未来，形状多样、大小各异的自动化运输车辆将会把商品快递到消费者的家，减少了运送到家的成本。如今的运送车辆必须考虑一个事实，由于车内有人类司机，所以车体的设计必须照顾到司机的安全与舒适。例如，运送比萨的车辆重量超过1吨，而比萨饼的重量只有几磅。相反自动运输车辆不需要气囊，或者沉重的、耐碰撞的框架，也不需要备胎、仪表盘

或是空调。车子结构紧凑，重量轻盈，造价低廉。

线上购物已经成功入侵了实体零售商店，每年人们在线上购买商品的数量都在持续增长。调研公司NPD集团通过调查全球165,000家商店发现，实体店的“购买访问”次数逐渐下降。“购买访问”是指消费者离开商店时，手里拿着买下的商品出门。在2012-2014年间，购买访问下降了13%，同时线上购买访问增长了21%。虽然实体店的购买访问仍要比线上购买访问在普遍性上多出四倍，但最终线上购物将占据个人购物趋势的主流。

随着亚马逊等大型零售商持续努力提升货运速度，改进货运方法，线上购物的趋势可能会保持增长。线上购物的吸引力之一是配送的便利性。大部分人做购买决策时考虑3C因素：成本（cost）、产品信任度（confidence）和便捷性（convenience）。传统上来说，消费者最终只能顾及其中两点。方便购买的、高质量的货物通常不便宜，而容易买到的平价货物质量一般。要找到价廉物美的货物并不容易。无人驾驶配送车辆能使线上购物更具魅力，并能通过快速的货运，让价廉物美的货物大量出现。一些零售企业将充分利用无人驾驶车厢成本低的优势，把整个实体商品的货品运送给消费者。例如，将来有个顾客想要买一双新鞋子，她会在线上预定几双候选的鞋子。那么无人驾驶汽车将会把这几双鞋子运送到她家，她试穿鞋子的尺寸与舒适度后，让车子把不合适的鞋子送回商店。

无人驾驶汽车多样化的造型能开创电子商务的新时代。今天，线上购物企业之间竞争的是谁能更快地配送货物。在密集的都市区域，当天抵达，甚至两小时抵达，正变得越来越普遍。快速而省油的车厢能“非常快”地把货物送到，这样的吸引力进一步破坏了实体店的优势。

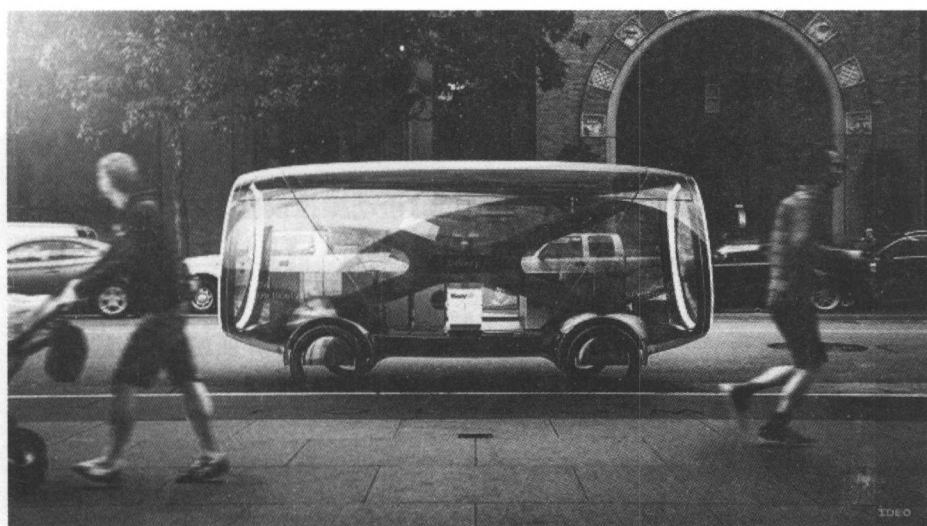


图11.5 自动配送卡车

来源：IDEO设计公司

犯罪与（成人）娱乐

任何对一项新技术的完整分析都不能回避其阴暗面——犯罪活动。计算机容易被用作盗窃数据和入侵自动银行系统窃取身份等的犯罪工具，无人驾驶汽车同样也吸引了跃跃欲试的罪犯将其应用在高科技犯罪中。

有些黑客会把他们的才能用于窃取和损害无人驾驶车辆的数据，包括传感器、数字地图和操作系统的数据；另一类型的攻击是劫持机器人，或是当无人驾驶汽车停在十字路口的时候，走到车前劫持车辆。对机器人进行劫持的犯罪行为，将可能成为今天某些绑架盛行的国家的严重社会问题。

劫匪将会利用许多无人驾驶汽车在编程中预设了“无论何种情况下，

优先保护人类性命”的特点，钻“安全特性”的空子，在车道上拦截无人驾驶汽车。如果一帮劫匪前后包围了无人驾驶汽车，那么车辆将会拒绝移动。大部分的劫车行为将会选择大货车，但也一定有会乘客车辆被劫持。因为没有方向盘和刹车，车内缺乏可用的手动操控设备，乘客便如同待宰的羔羊，陷入了一场可怕的噩梦。

另一种可能产生的恶作剧是，人类司机忽远忽近地驶向无人驾驶汽车。当人类以“Z”字形行驶在高速公路上，来回穿梭于无人驾驶汽车附近时，将会制造出一片恐慌。然而，这类破坏行为会渐渐消亡。不过当车辆的机器学习软件掌握了这类飘忽不定的驾驶规律后，恶作剧制造者可能又会想出其他新的难以预测的把戏。

说到潜在的恶作剧与难以预计的行为，当人类不再操心驾驶的时候，什么可以规范人类在无人驾驶车内的行为呢？根据卡内基·梅隆大学的调查，大部分的乘客会在车内使用移动设备。名单上的行为接下来是吃东西、阅读、看电影和工作。一旦人们不用再操心驾驶，他们就能找到新的娱乐方式。

今天，大部分司机在驾驶时最常见的娱乐行为就是听广播。事实上，调查显示人们收听广播的接近半数时间是在车内。当人们不再局限于在车内听广播，广播的收听率将会直接下降，而电影和电子游戏的使用率将直线上升。

今天车内简陋而设计差劲的车载娱乐系统将不复存在。豪华的移动办公室或个人通勤车将配备娱乐设施，车厢无关紧要的装置将被拆除。乘客将把时间花在在娱乐设备上或者利用车载网络进行在线社交。

对于少数拥有无人驾驶豪车的人来说，过去花在驾驶上的时间也可能被用来进行某些不良嗜好：性行为、毒品或酗酒。我们可以想象未来人们有可能放肆地沉溺在这些恶习中。每次当我们谈及无人驾驶汽车的时候，随即便有人悄悄靠近我们询问是否想起以上这三种禁忌话题，每个人都能马上联想到但不会公开谈论的话题。然而在封闭式车厢内，我们可以大声承认，因为车厢保护个人隐私，人们会在车内大肆进行这三种罪恶的活动。

事实上，我们对这类活动有明确的规定，至少在车内吸毒和酗酒属于违法的行为。如果在无人驾驶汽车内部装配上摄像头，一些有趣的事情将会发生。深度学习软件将学会检测车内的吸毒或酗酒的行为。如果政府努力地想要弥补原本在超速和违规停车的罚款收入，那么市政人员便会要求车辆上报车内乘客的不当行为。但是，在无人驾驶汽车中的性行为很可能不算违法（只要是在两名成年人之间自愿发生，且不涉及金钱交易）。所以一些新型的无人驾驶汽车可能沦为安装有遮蔽窗户的“派对巴士”。

色情产品的消费一直对科技技术的发展起了助推的作用，例如VHS家用录像机流媒体和互联网。无人驾驶汽车将成为色情产品用户放荡其中的新“乐园”，尤其是当Oculus Rift之类的虚拟现实头戴显示器等产品普及之后。

未来之路

未来前景如何？机器人技术已发展到了重要的转折点，无人驾驶汽车

如期成了安全实用的运输方式。我有时候在想，如何向下一代解释，驾驶汽车曾经是成年与自由的象征。当我向他们讲述高中时期的我和同学们曾经用一个学期的时间来学习驾驶技术，我能想象出他们的脸上写满不屑和难以置信的神情。

我期待着有一天，虽然我们都老去，却仍然记得亲自驾驶车辆的感觉。就像文字处理器通用的今天，人们会怀念使用老式打字机的日子。智能化自动汽车不久便会以高效、舒适的体验俘获我们的心。以至我有些怀疑，我们当中是否还有人真的怀念“手握方向盘，脚踩刹车”的日子，这些对“驾驶的麻烦工夫”的记忆都被蒙上了一层浓浓的怀旧色彩。

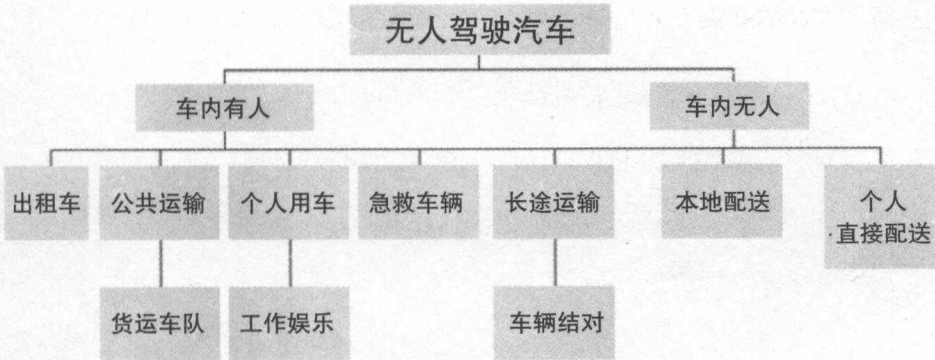
也许，人们会想出一种新的娱乐方式：在练习场体验手动驾驶。未来的练习场并不是用来打高尔夫球，而是驾驶着老式汽车兜圈。这是一种老年人的休闲方式，而年轻人也能在围观中感受到一丝嘲讽的快乐。我将会是这些老年人的其中一员。

几十年后我步入退休生活，发觉剩余的人生屈指可数，想必我会大吃一惊并果断加入练车场的会员。让我们暂且把它称之为“悠驾（U-drive）”吧。就像一位有个性的作家在老式打字机上，行云流水般流畅地击打着键盘，我时而渴望速度的快感，时而在堵车的车队中望着方向盘发呆。不过，还是先让我戴上保护头盔，召唤无人驾驶汽车载我抵达练习场吧。

到达场地后，我将冲向服务台，希望能驾驶急救车或警车。在赛道上，没有什么可以比在警笛呼啸中的高速驾驶更帅了！我转动着方向盘，就像回家一般亲切，不禁随着车内的古董收音机放声高歌！同时，我会过

瘾地按响喇叭，直到引起其他司机的不悦才恋恋不舍地停手。

将来的某天，我定会摘下乘坐无人驾驶汽车需要的保护头盔，去体验手动驾驶的乐趣。如果我还没有忘记如何熟练驾驶的话，我甚至不想系上安全带，还要将油门踩到底，不去理会道路两旁的限速标志。不过，在未来的驾驶场景中有一样事情可以确定：“悠驾”公司并不愚昧，他们才不会告诉用户，在即将发生交通事故时，这些老式汽车将自动且迅速地切换回无人驾驶模式，方向盘和油门瞬间失效！



重要属性

按需运输	共享班车	移动办公	优先级最高	长途大量运输	批发配送	小型
短途运输	自适应线路	长距离		节省能源	优先级低	优先级高
卸货即走	多功能	高速			多停靠点	电子商务配送
高实用性	长途车辆结对	全天候			农产品直供	
		平稳				
		奢华				

图11.6 无人驾驶汽车分类

后记

寒武纪大爆发

Afterword:
The Cambrian
Explosion

寒武纪大爆发是生物进化史上的重大事件，在此之前绝大部分主流的动物门尚未形成，只存在简单的生命形态；在此后的几百万年中，生物多样性加速发展——开始出现多细胞机体，生命形态变得与我们今天所熟悉的动物王国非常类似。其中的某些多细胞动物，如威瓦西虫（*Wiwaxia*）和五个眼睛的欧巴宾海蝎（*Opabinia*），长得一点也不像我们认识的生物，倒是像极了以外星人为主题的恐怖电影中的生物。不过，产自那个时代的大多数化石都有今天地球上多种多样的生命形态的先祖影子。

史蒂芬·杰伊·古尔德（Steven Jay Gould）在其1989年的著作《精彩的生命》（*Wonderful Life*）中，推广了寒武纪大爆发的观念并探索了它的起源之谜。为什么所有的现代生物都出现在一个相对短暂的时期？突如其来的生物多样化出现，貌似与达尔文进化论认为的“逐渐持续的进化过程”相矛盾。“达尔文在《物种起源》（*The Origin of Species*）中提到，生物种类没有任何先兆，突然出现，毫无疑问，寒武纪大爆发严重地违背了他物竞

天择的理论。”

寒武纪大爆发之谜使得史蒂芬·杰伊·古尔德，以及同时代的哈里·惠廷顿（Harry Whittington）和尼尔斯·埃尔德里奇（Niles Eldredge）用一种现代观点来看待进化论，那就是间断平衡理论（Punctuated equilibrium）。根据该理论，进化过程是由近乎停滞的渐变微进化与跃变式的大进化交替出现组成的。

机器人的进化

确实，当我们进行机器人进化实验时，整个过程非常不稳定。进化机器人学说（Evolutionary Robotics）是机器人研究领域的一个分支。该学说应用于好几代机器人的变异和选择，刺激了计算机朝“达尔文理论”的进化特点发展。在实验中，我们让计算机随机地把机器人零部件（机械关节、刚性构件、发动机、电线和神经元）放在一起制造虚拟机。然后对计算机编程，使其快速选择并改良组装好这些零件，制作出“下一代机器人”，再把这些“下一代”纳入机器人队伍，再循环进行实验。我们则站在一旁观察。

第1代机器人全部都是一堆零件，而且都不能移动。当实验进行到至少第100代机器人后，那些零件依然不能移动，看起来我们不会在这项实验收获任何有趣的发现。进行到大约120代，那些电线、发动机、关节和连接看上去依旧是随机安排的，但它们突然能够产生小幅度的颤动或移动。尽

管动作幅度微小，却绝对要比另一堆丝毫不动的零件要好得多。于是，这些动态机器人便在虚拟世界渐渐出现。回想起来，我们把这种现象命名为“震动发现（discovery of vibration）”，在整个机器人世界里，造成了突然的间断的性能改进。几百代之后，又有了新的发现，机器人性能再次得到明显改进。实验继续。又进化了几百代人后，我们看到工作性能优异的机器人，它们中的部分形态怪异如外星生物，另一些则能依稀辨识出熟悉的形态。

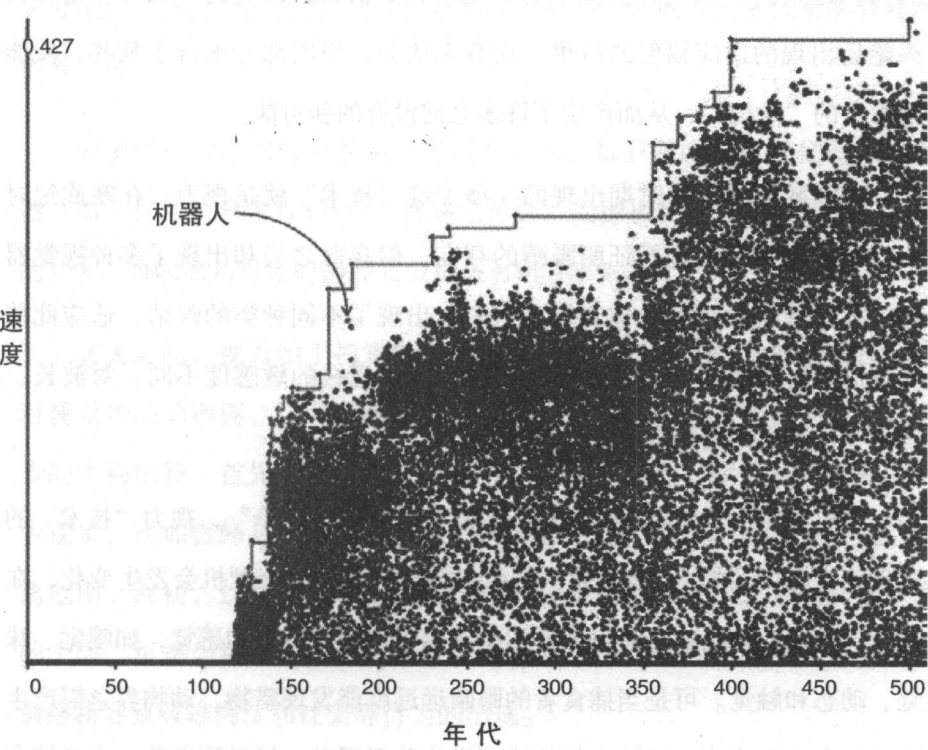


图12.1 数百代机器人的进化过程

传感器王后

实验中，机器人的人为进化最终得出满足的结果，然而生物界进化过程却鲜有规律。寒武纪大爆发的成因到今天依然是个未解之谜。有人把这史前出现的多样性形成原因，归为多细胞生物发现成果的一部分。也有人认为，这是因为氧气等资源丰富、地球的生物宜居性提高，或者在物种大灭绝后出现的适应辐射的后果。还有人认为，原因部分来自于某些“使能技术”的“发现”，从而产生了许多之前没有的新可能。

在寒武纪大爆发时期出现的一项关键“技术”就是视力。在寒武纪时代以前，没有任何化石证明眼睛的存在，但在此之后却出现了多种视觉器官。在寒武纪中后期的伯吉斯页岩中，出现了不同种类的眼睛，适应此前出现的多种器官形态——视觉敏锐度不同，光线的敏感度不同，对波长、动态和色彩不同的识别能力。

我们在第一章中提到安德鲁·帕克的“光开关理论”，视力“技术”的出现使生物的“捕食-被捕食”关系中的对抗性以及交配机会发生变化；在视力出现以前，捕猎与逃生依靠的是作用于近距离范围的感觉，如嗅觉、味觉、动态和触觉。可是当捕食者的眼睛远远就能发现猎物，动物界急需产生新的捕食攻防战略，导致生物的协同进化防卫展开一场激烈竞争。由于捕食者视力的提高，猎物必须学会如何隐藏自己，如何快速逃生，如何用迷彩色伪装或长出刺来，这些“新技术”带来生物形态的进化。

我们无法确切知道视力是否在寒武纪的生物多样性出现中扮演了重要角色，但我们可以在此做一个假设：并非眼睛本身带来了多样性，而是其后的认知能力的发展。区别于触觉、味觉和嗅觉，视觉信息是“高带宽”的，在空间分辨率和时间流中，导致数据率比其他感觉器官要高。因为数据是远距离的传感器，范围覆盖机器以外的广阔世界，需求新的场景分割认知器官，空间模型和对世界的了解。也许数据大量地在早期的大脑中蔓延，是因为视觉为个体生物提供了稍微好一些的认知能力。认知能力的发展带来许多新机会：新的“捕食-被捕食”策略、新的寻找配偶策略和新的资源发现。

提供视觉信息的机制控制了我们的脑。每只眼睛含有15亿的光传感器，而一只耳朵只含有3万个声音感觉神经。处理视觉信息的神经元占皮层的30%，而触觉和听觉的神经元分别只有8%和3%。

毫无疑问，视力始于视觉器官与视觉皮质协同进化的。神经器官包括对视觉场景的解读，以及各种很快被发现的“新应用”，形成了阶段性进步的生物创新。首先，视力对于协同进化作出了贡献，像是器官之间的共生现象，比如蜜蜂与开花植物之间的关系。它也扩大了个体寻找配偶的距离范围。起初，这种先进的寻找配偶感知功能也许只是用来发现同一种类的个体。可是当视觉进一步提升时，更有利于分辨配偶的条件是否理想，最终将导致雌雄淘汰和社交等行为的出现。

生物界的寒武纪大爆发与机器人的前景探索具有极高的相似性。DARPA计算机项目的前负责人吉尔·普拉特（Gill Pratt），在2015年被任命为丰田公司自动驾驶汽车分部的领导时写道：

今天，好几项前沿技术的发展正在机器人的多样性和能力发展范畴燃起一场大爆发。机器人依赖的许多基础硬件技术，尤其是计算能力、数据储存和交互能力，正在以幂次数量级的速度增长。两项新兴技术——云计算和深度学习将会影响上述基础技术在良性循环中进行爆炸性发展。

确实，对机器人影响重大的基础技术正飞速发展，这些技术使自动驾驶汽车的潜力和多样性成为可能。

1. 动力储存与效率的幂次增长。

自动机器人需要动力的自动化；在过去几十年中，电池技术不断进步。从1950年的铅酸电池到今天的聚合物锂离子电池效能提高了两倍。除了电池能力以外，机器人甚至能从效率的提高中获得更大的动力，例如发动机效率的提升。动力储存和效率的提升加速了自动系统整体性能的提高。性能更优的机器人可以花更多时间做事情和学习，降低充电和获取动力的时间。

2. 计算能力的幂次增长。

正如摩尔定律所预测的，每一块钱所能购买的可用计算能力每18个月增加一倍。由于受到体积限制，最近几年的晶体管小型化发展速度下降，然而每单位美元所能购买的计算能力却继续以其他方式来增长，如多核平行等。对于需要处理数据流和实时决策的自动系统而言，计算能力真是至关重要。高速的处理能力能让机器人在结构化程度较低的环境中顺利运行，更快地从经验中学习。

3. 传感技术的幂次增长。

从激光雷达发展到超声波，传感器变得更精确，速度更快，价格更低。所有维度中快速发展的传感器之一是相机。受到移动设备发展驱动，相机技术的性能和价钱都在以幂次数量级发展。光学仪器和传感器在成本、大小、性能的发展使得多功能的相机可以媲美一个机器人。多重数据流带来更好的认知性能，因为人们可以从多个视点（例如，来自超立体视觉的深度感知和速度感知）获得更可靠的场景理解，同时在应对恶性破坏和暂时的传感盲区时，带来更高的鲁棒性（robustness）。

4. 数据储存的幂次增长。

数据储存能力正在以幂次数量级的速度不断增长。这些提高影响的不仅仅是每一块钱能购买多少储存字节，还影响了数据存储和检索的速度和可靠性，每次数据存储消耗的能量，数据的物理重量（每千克的字节数量）。当机器人能高效地在本地储存大量数据，它们能回忆并重复利用以前的经验，从过去储存的经历中学习出新的知识。

5. 交互带宽的幂次增长。

短距离和长距离的带宽都在过去几十年内得到了迅猛增长。几十年前，信息传送不仅速度慢、技术困难、价格昂贵，而且效果不佳。今天我们通过卫星传输以太字节（Terabyte）为单位的信息量，丝毫不会怀疑信息是否能完整到达目的地。长距离的交互能力与可靠程度，以及允许机器人相互之间分享数据和本地结果分析，催生了联合分享智能技术，就是所谓的云机器人（cloud robotics）。

幂次增长之王——算法

人们往往倾向于留意到计算、交互、传感方面的机器人硬件技术的飞速进步，但我们常常遗忘是哪些发现带来了这些进步及其背后的新算法技术的发明。

在计算机科学家和电子工程师中流传着一种说法，无论硬件工程师取得了什么样的进步，软件工程师是都能马上“废了”它们（原本程序员的说法用词太过激烈，不适合写在书里）。我们都知道无论计算机的硬件性能变得多快，操作系统软件却总是运行太慢，无法跟上速度。然而事实刚好相反。

算法的进步并没有类似“处理速度/年”“兆像素/美元”等单位，一直以来都难以量化。因为算法在执行各项任务时形态各异，目标也不断变化。举一个例子，我们用算法来解微分方程。这一类的数学算法对于任何需要进行预测和动态控制的机器人来说，都是关键的因素。1945–1985年间，执行这类基础任务的算法以每年3万点的系数在增长，或者说是年均29%的增长率。该增长率达到了同期的基础硬件的平均水平。

另一个例子是数据分析的算法进步。经典的快速傅里叶变换算法（fast Fourier transform, FFT）被用于几乎所有的数字信号处理系统，比起原始算法，这一算法更凸显了指令的速度提升的重要性。但究竟FFT算法带来了多少增速呢？事实证明进步因素很大程度上取决于被分析的数据库大小。若是小型的数据库进步空间小，大型的数据库进步程度相当大，这种进步

甚至需要硬件用几十年的发展时间才能跟上它的步伐。正是因为算法的进步，我们才有了今天数据分析的方法，否则即使摩尔定律继续作用100年，这种方法也不可行。

不像硬件的进步曲线那么平整，算法的进步更像带有“间隔平衡”的特点，或据推测，它像是进化中的系统。算法的进步不是一帆风顺，而是断断续续地前进。好比一个生态系统，算法的进步，是在与众多其他算法的竞争中诞生的。某些算法诞生后，因其学术性深奥难懂而被埋没；其他算法在发展壮大后不久随即消亡，因为有更好的算法占据了市场，或是它们能解决的问题不再受到重视。一套算法需要经过不同人的多次修改，直到偶然的机会它拥有了广泛的识别能力。

人工智能算法也不例外。在一个多世纪的艰难前行中，人工智能算法不断进步又遭遇寒冬，风光一时却也曾经黯然失色。只是无论处于领导地位的是哪个阵营的人工智能，在过去几十年中人工智能算法总体上还是成绩斐然。我们知道，无论处理器的速度如何飞快，数据储存或相机分辨率如何发展，都不能使罗森布拉特的感知器稳定地分辨出猫和狗的差别——即使再强大的计算能力也不能使20世纪90年代的标准双层神经网络顺利完成这个认知任务。只有在大规模视觉识别挑战赛上的竞争激烈的各种算法中才能找到某种优秀的算法来完成识别任务。就像是岩石缝中生存的哺乳动物，卷积神经网络最终打败了人工智能的传统强劲对手。

当深度学习算法在断断续续的进化过程实现了阶梯飞跃，整个人工智能界沸腾了。当该行业的重心从硬件转移到软件上时，自动行业也感受到了类似的震撼。

算法的进步大量涌现

我们不知道眼睛在大脑的发展过程中起了何种促进作用，但我们知道智力的作用远远超过视觉皮层细胞，从识别捕食者和猎物、寻找配偶，到今天阅读本书时用到的充分交互与自我认识能力。我们知道深度学习算法从最初的为视觉感知服务，到今天成功应用到许多人工智能领域，例如语音识别到语言产生再到艺术创作。我们推测这股发展趋势仍将继续。

发展趋势将持续多久？何时结束？这些都是科幻作家和哲学家考虑的问题。如果我们用硬件能力为基础来推测人工智能的发展，发展过程将持续到2020年，到那时计算机能力将于大脑的计算能力相当。可是预言似乎不能满足我们。我们想知道什么时候计算机能拥有与人类水平相当的智能行为方式，而不仅是它们原始计算能力的发展。问题在于这类设想要走向现实可谓难上加难。

计算机要发展出类似自我觉醒或自我认知的能力，需要的不仅仅是硬件的提升，而是一套全新的算法。况且我们不能确切知道“机器的自我认知”是什么，我们只知道这不用于下棋或开车，它没有明显的规律可循。因此不像许多科幻电影描述的场景，这不是由软件开发的高手通过编程可以直接解决的问题。

相反地，机器的自我觉醒发展缓慢。自我觉醒是什么？我们采用实际的定义：它不过是一种自我刺激的能力，用来预测现在行为所带来的后

果，从而不需要在现实中真的做出该行为。你能想象出明天的你行走在沙滩上的情形吗？你能想象出大海的气息和踩着沙子的感觉吗？感觉很美好吧，让你心动真的要行动了吗？如果答案是肯定，那么你进行的就是“自我觉醒”。有人会说，就连能引起感知的情绪，比如恐惧和欢乐，都能在过去的经历基础上，把未来的行为结构投射到我们目前的状态中。比如，“痛苦”象征着现在遭遇的破坏，而“恐惧”表示极有可能发生的严重破坏，然而“担忧”在内在自我模式的基础上，反映出严重性轻微、距离更长远的负面结果。

如果机器人根据目前的行为预测到将来的场景，然后开启预测模式计划将来的行为。在某种程度上说，它自我觉醒了。2006年我们展示了一款能建造自我形象的机器人——虽然图片描绘并不准确，呈现是粗糙的类似剪贴画的小人像——但机器人却能从此离开物理实验和外部编程来学习行走。可是当时机器人制作出自画像时，感知器和预测算法已经达到了极限。

也许当深度学习算法应用到所有人工智能领域时，我们才能看到新一代的机器人画出越来越准确的自我肖像以及周围环境的模式，这样渐渐发展成自我意识觉醒。

一辆自我意识觉醒的汽车不会在车道上遇到你，或和你聊聊路况，也不会很贴心地顾及你感兴趣的事务。但一辆自我意识觉醒的汽车能够越来越准确地建造于驾驶模型，以及你的驾驶习惯——哪些它可以做到，而哪些在它能力范围之外，这些潜在动作的风险利弊可能是什么，等等。就像我们人类的自我意识觉醒，能够超越自身范围，领会他人的感受和意图，汽车也能明白路上行驶的其他车辆的下一步举动。

我们估计，从多细胞进化出视觉感知细胞，花了500万年；从视觉感知细胞进化到智人花了5亿年。如果依照这个时间做类比，感知器花了50年时间从20世纪50年代的“无视力机器人”发展而来，也许要花上500年的时间，人工智能才能发展到与人类水平相当自我意识觉醒。硬件进化能力这股发展趋势，可是算法的进化必须遵循其自身的间隔发展特点。

不过，无论是2020年还是2500年，相比于人类的进化历史来说，都太快了。

人类一直以来都倾心追求制造出有生命的活物，古时候的炼金术士用尽各种办法，赋予泥石生命的气息。历史上出现了各种神秘的处方，多年以后，炼金术士被他们的现代“追随者”——机器人学家所取代。今天，研究机器人的专家拥有更好的工具、更深层的理解力和稍微充裕一些的资金。最终，我们依然在不停尝试，把生命气息带给毫无生命特征的机器。

致谢

几年的时间里，我们与众多相关人士进行了很多激动人心的、发人深思的讨论，最终才形成了这本系统性介绍的书籍。我们要特别感谢所有在本书成书过程中伸出援手的人，他们提供的任何形式的帮助都有巨大作用。

我们想要感谢在自动驾驶汽车研究领域长期作出突出贡献的两位学者，他们是奇点大学的布拉德·邓普顿（Brad Templeton）和普林斯顿大学的阿兰·肯豪森（Alain Kornhauser）。想要找到专门在机器人领域有着深厚研究并且应用到汽车领域的专家并不容易，这两位学者不仅知识渊博，并且还愿意无私地与我们分享他们的真知灼见。通过与布拉德、阿兰交谈、阅读他们的博客和与他们邮件往来，我们对无人驾驶汽车领域有了更多深刻的认识。

我们还要感谢卡内基·梅隆大学国家机器人工程中心（National Robotic Engineering Center, NREC）的全体成员，尤其是布莱恩·扎伊克（Brian Zajac），他曾花费数小时带领我们参观了他们的设备和所有自动化的产

品。感谢阿姆农·沙舒亚博士（Dr. Amnon Shashua），作为“移动眼”公司的创始人，他无私地向我们提供了有关计算机视觉方面的材料。

康奈尔大学和哥伦比亚大学的创新机器实验室（Creative Machines Research Lab）的历届学生在自动系统和人工智能方面都作出了诸多贡献，没有他们的智慧付出，我们的工作也无法开展。尤其是杰森·奥辛斯基（Jason Yosinski），他是认识到“深度学习”的价值的先驱之一。正是由于杰森的坚持，我们才能数年如一日地开发深度学习软件，最终使得深度人工神经网络系统如今被广泛接受，对此我们表示由衷的感谢。

我们也要感谢麻省理工学院出版社团队的辛勤工作和无私奉献，尤其感谢玛丽·勒夫金·李（Marie Lufkin Lee）、米歇尔·西姆斯（Michael Sims）以及凯思林·亨斯里（Kathleen Hensley）对本书的热心支持。

最后，我们还要感谢那些曾经在幕后工作的工程师、发明家、创新者、艺术家和企业家，如果没有这些富有创造力的、大胆的创新者的奇思妙想，也就不可能会有关于无人驾驶的书籍。感谢你们的付出，相信在不久的将来，世界上各个街道、停车场以及高速公路将成为更加安全、美好之地。

图书在版编目 (CIP) 数据

无人驾驶 / (美) 胡迪·利普森, (美) 梅尔芭·库曼著; 林露茵, 金阳译. — 上海: 文汇出版社, 2017. 5

ISBN 978-7-5496-2100-2

I. ①无… II. ①胡… ②梅… ③林… ④金… III.

①无人驾驶—汽车 IV. ①U469.79

中国版本图书馆CIP数据核字 (2017) 第090453号

Driverless: Intelligent Cars and the Road Ahead by Hod Lipson, Melba Kurman

Copyright © 2016 by Hod Lipson, Melba Kurman

Published by arrangement with MIT Press

Simplified Chinese translation Copyright © 2017 by Shanghai Dook Publishing Co., Ltd.

All right reserved.

版权登记号 图字: 09-2017-243

无人驾驶

作 者 / 【美】胡迪·利普森 【美】梅尔芭·库曼

译 者 / 林露茵 金 阳

责任编辑 / 刘 刚

特邀编辑 / 姜一鸣 姚红成

封面装帧 / 陈艳丽

出版发行 / 文汇出版社

上海市威海路 755 号

(邮政编码 200041)

经 销 / 全国新华书店

印刷装订 / 三河市龙大印装有限公司

版 次 / 2017 年 5 月第 1 版

印 次 / 2017 年 5 月第 1 次印刷

开 本 / 710mm × 1000mm 1/16

字 数 / 245 千字

印 张 / 21.75

ISBN 978-7-5496-2100-2

定 价 / 58.00 元

侵权必究

装订质量问题, 请致电010-85866447 (免费更换, 邮寄到付)

胡迪·利普森

Hod Lipson

人工智能、机器人学专家，哥伦比亚大学人工智能实验室主任，原康奈尔大学人工智能研究院负责人，著有《3D 打印》一书。利普森教授著有 200 余篇论文，被各大学术刊物引用超过一万次。

梅尔芭·库曼

Melba Kurman

在微软和康奈尔大学从事人工智能实验和产品研发超过 15 年，现在是一家创新公司的总裁，著有《3D 打印》一书。

· 读客® 商业类图书 ·

相关推荐

《分享经济的爆发》：全球分享经济泰斗揭示分享经济将从哪些方面重构我们的社会。

《爆品战略》：39 个超级爆品案例的故事、逻辑与方法》：传统企业转型、互联网创业的实战指南！爆品是一种极端的意志力，是一种信仰，是整个企业运转的灵魂！

《摩根财团：美国一代银行王朝和现代金融业的崛起（1838~1990）》：亚投行首任行长金立群“历时二载，凡三校”呕心校译。关于华尔街，读这本就够！

请关注新浪微博：@读客图书
每天都在免费送超级畅销书！

策 划：读客图书

版 权：读客图书 021-33608311

责任编辑：刘刚

特邀编辑：姜一鸣 姚红成

封面设计：陈艳丽

网 址：www.dookbook.com

投稿邮箱：dookchina@163.com

招聘邮箱：hr@dookbook.com

无人驾驶

你初次拥有的智能机器人，很可能就是你的汽车，
你会把自己的性命交给它！

无人驾驶汽车将成为人类可以放心托付生命的初代自主式智能机器人。伴随着无人驾驶技术的普及，人工智能会彻底颠覆人类对时间和空间的认知，重塑人类社会的出行、居住和商业结构。

一场关于交通、物流、能源、制造、保险、医疗、伦理道德的全面挑战已经开始。人类如何实现向人工智能时代平稳过渡？

哥伦比亚大学人工智能实验室主任利普森教授
将全面解读人工智能时代的第一次技术变革。

任何对无人驾驶汽车的技术、社会影响和人工智能未来发展感兴趣的人都不能错过的伟大作品。这本横跨多领域和多学科的书，为无人驾驶这项技术创造了全景式的阅读享受。

—— 丹妮拉·鲁斯 (Daniela Rus)

麻省理工学院计算机和 AI 实验室主任

我们预计无人驾驶车辆将全面接管我们的道路和生活，所以任何人都需要对这项颠覆性技术的影响和我们所面对的问题作整体性的评估和思考。这本书是帮助我们实现最平稳过渡的伟大路线图。

—— 史蒂文·贝肯菲德 (Steven Berkenfeld)

巴克莱银行董事总经理

建议上架：人工智能/计算机
ISBN 978-7-5496-2100-2



9 787549 621002 > 定价：58.00元

发扬传统文化 成为文化传统
www.dookbook.com

